objection to the presence of non-elected subject matter. Basis for the amendment of Claim 13 may be found, *e.g.*, at page 5, lines 10-20; at page 15, lines 3-16; at page 31, lines 22-26; and at page 32, line 7 to page 33, line 11. of the specification. No new matter is added.

It is noted that a continuation application was filed on April 23, 2003; the serial number is not yet available.

The specification and abstract are amended to correct obvious typographical and spelling errors. The amendment to the paragraph beginning on page 89, line 30, through page 90, line 10, of the specification replaces "SEQ ID NO: 36" with —SEQ ID NO: 47—. The amendment finds basis in the sequence listing, in which SEQ ID NO. 47 is the amino acid sequence of the VP16 minimal domain trimer. No new matter has been added.

Included as an attachment is a marked-up version of the specification paragraphs and claims that are amended, per 37 CFR §1.121. A Supplemental Information Disclosure Statement is also filed on the same day herewith under separate cover.

RESPONSE TO APPLICANT'S REQUEST FOR CLARIFICATION OF THE RESTRICTION REQUIREMENT

Responsive to Applicant's request for clarification of the Restriction Requirement of September 11, 2001, it is maintained that SEQ ID. NOS. 1-18, set forth as exemplary fusion proteins in the instant application, are restrictable as "independent and distinct inventions" because "sequences encoding different proteins are structurally distinct compounds and are unrelated to one another." The Examiner asserts that the generic claims have been examined, as well as SEQ ID. NO. 1. It is further alleged that assuming, *arguendo*, that the generic claims are allowable, SEQ ID. NOS. 2-18 will not be rejoined because they are not elected species but constitute "independent and distinct inventions."

It is respectfully submitted that by restricting the claims to a single nucleotide sequence, genus claims such as, for example, Claims 1 and 20, are not examined. The pending claims are drawn to the generic concepts set forth,

e.g., in Claim 1, a fusion protein that is a ligand activated transcriptional regulator containing a nucleotide binding domain containing modular zinc-finger peptide units operatively linked to a ligand binding domain from an intracellular receptor; and Claim 20, which has the additional element of a transcription regulating domain. The Examiner states that these generic claims, along with the remaining pending claims, have been restricted to a fusion protein encoded by a single nucleotide sequence, namely, SEQ ID. NO. 1.

As discussed previously, Applicant respectfully submits that if the claims are restricted to a single nucleotide sequence, then such a restriction would be improper as applied to the instant claims.

The claims are directed to fusion proteins that are ligand activated transcriptional regulators, which include a modified ligand binding domain. This is a generic product and a variety of components can be used. Absent a finding of a art that describes a species of such construct, there is no reason to limit it to a single sequence of nucleic acids. Countless examples of each component of the fusion protein are known to those of skill in the art. It is the instant application that teaches a construct that contains such components. In view of the instant disclosure, one of skill in the art can readily identify the components and prepare fusion proteins that contain the requisite elements. These claims do not fall under the rules that govern the examination of applications directed to nucleic acid molecules. The searchable aspect is not the specific sequences but the fusion as a fusion of particular functional elements.

The genus claims, *e.g.*, Claims 1 and 20 are linking claims, and as such must be examined with elected species. These claims "link" the dependent claims, which specify further limitations of the fusion protein domains and set forth exemplary nucleotide sequences (*e.g.*, SEQ ID. NOS. 1-18) encoding fusion proteins having the structural and functional elements of the genus claims. According to MPEP §809, when claims linking more than one group are found, the Restriction Requirement must be conditioned on 1) specifying the linking claims; and 2) examining the linking claims with the elected group. The

linking claims **must** be examined with the elected group, and the Restriction Requirement **must** be conditioned on allowability of the linking claims. If the linking claims are deemed allowable, then the Restriction Requirement must be withdrawn and all claims directed to nonelected subject matter which depends from or includes all the limitations of the linking claims must be rejoined.

In this instance, by restricting the instant claims to a single nucleotide sequenc and requiribng amendment of the claims to exclude the genus, the Examiner has improperly excluded the generic linking claims from the Restriction Requirement. Moreover, it is respectfully submitted if the Examiner considers the claims as being drawn to SEQ ID. NO. 1, which contains a zinc finger peptide sequence linked to an estrogen receptor ligand binding domain, then the Election of Species requirement, responsive to which Applicant elected the species "estrogen receptor", has been rendered meaningless.

Inconsistencies in the Requirement for Restriction and Election of Species

Applicant respectfully submits that it is contradictory to have dependent claims (*e.g.*, claims 9, 10, 13, 18, 19) drawn to variants of a generic fusion protein, yet restrict the generic fusion protein (such as that set forth in, *e.g.*, Claims 1 and 20) to one specific molecule having a particular sequence (*i.e.*, SEQ ID NO. 1). Further, as discussed below, the Election of Species requirement is rendered meaningless and improper if the claims are restricted to a single nucleic acid sequence.

In addition to the inconsistencies noted above that make the Requirement for Restriction as set forth improper, restriction of the claims to a single sequence, if such was intended, is improper in the instant context. Restrictions to single nucleotide sequences is discussed in §803.04 of the Manual of Patent Examining Procedure (MPEP). According to MPEP §803.04, claims drawn to nucleotide sequences encoding different proteins are deemed properly restrictable, although the Commissioner has decided *sua sponte* to partially waive this requirement for a reasonable number (usually, ten) of patentably distinct sequences. The restriction of nucleotide sequences applies to

"applications **claiming** more than ten [usually] individual independent and distinct nucleotide sequences" (MPEP §803.04; emphasis added).

The instant generic claims are not drawn to particular nucleic acid or protein sequences, *i.e.*, particular sequences are not claimed. Rather, they are drawn to a fusion protein with specified structural and functional limitations that can be searched without the selection of particular sequences. Thus, MPEP §803.04 does not apply to the instant claims. SEQ ID NO. 1 is merely an exemplary molecule having the structural and functional elements of the generic concept set forth in Claim 1.

The structural and functional elements set forth in the generic claim can be searched by virtue of their properties, *e.g.*, ligand-binding domains of intracellular receptors and zinc finger proteins. By requiring restriction to a single sequence, the Examiner, without citing any art, is urging that a fusion protein with components having the structural and functional properties as claimed is not patentable unless it is drawn to a particular fusion protein sequence. The Requirement for Restriction provides no evidence of record that the claimed fusion protein is only patentable as to particular sequences, nor are there any other reasons set forth for requiring restriction of the generic claim to particular nucleic acid sequences.

The Election of Species requirement has been rendered meaningless

Moreover, if the Examiner considers the claims as being drawn to a single nucleotide sequence, the restriction to SEQ ID. NO. 1, which contains a zinc finger peptide sequence linked to an estrogen receptor ligand binding domain, would render election of the estrogen receptor in response to the Election of Species requirement meaningless. The instant claims can only be restricted to the designated species, *i.e.*, estrogen receptor, if no generic claim is deemed allowable. If a generic claim is ultimately allowed, then the additional claims directed to specific elements within each category, written in dependent form or otherwise including all the limitations of the allowed generic claim(s), will necessarily be allowed. In the Requirement for Restriction, the Examiner has

restricted the generic claim to a fusion protein in which the intracellular receptor is the elected species, *i.e.*, the estrogen receptor (the sequence of the intracellular receptor set forth in SEQ ID NO. 1), without making any determinations as to allowability of the generic claim. Therefore, by restricting the generic claim to a particular sequence, *i.e.*, SEQ ID NO. 1, the Election of Species requirement is meaningless and improper.

In light of the above, Applicant respectfully maintains that the Requirement for Restriction as set forth in the instant application is improper.

OBJECTION TO CLAIM 25

Claim 25 is objected to as being drawn to non-elected subject matter. As discussed above, restriction of Claim 25 to a single sequence is improper because the generic claim is to a fusion protein with specified structural and functional domains that can be encoded by more than one sequence. Nonetheless, this objection has been rendered moot by amending Claim 25 to specify a fusion protein encoded by SEQ ID No. 1.

THE REJECTION OF CLAIMS 1-3, 5-35, 37-46 AND 69-73 UNDER 35 U.S.C. § 112, FIRST PARAGRAPH

The Enablement Rejection

Claims 1-3, 5-35, 37-46 and 69-73 are rejected under 35 U.S.C. § 112, first paragraph because it is alleged that the specification, while being enabling for a fusion protein containing an amino acid sequence set forth in SEQ ID NO. 1, and for fusion proteins containing a zinc finger protein E2C(Sp1), B3B(Sp1) or B3C2(Sp1) and ERD, KRAB, SID domains that bind the *erbB-2* promoter, does not reasonably provide enablement for a protein variant of SEQ ID NO. 1, or a fusion protein containing a ligand binding domain from an intracellular receptor where the nucleotide binding domain is from a zinc-finger peptide that binds a sequence of at least 3 nucleotides, or a variant of such a fusion protein. It is further alleged that the rejected claims are overly broad because insufficient guidance is provided as to which of the myriad of variant fusion polypeptides will retain the characteristics of functioning as a transcriptional regulator. The

Examiner cites Voet *et al.* (*Biochemistry*, John Wiley & Sons, (1990)), for the proposition that even single amino acid changes can have dramatic effects on a protein's function, structure or architecture. The Examiner concludes that while the claims are enabled for polydactyl zinc finger fusion proteins that bind the erbB2 promoter, they are not enabled for polydactyl zinc finger fusion proteins that bind to any other contiguous nucleotide sequences of at least 3 nucleotides, or a variant of such a fusion protein and it would allegedly require undue experimentation for one of skill in the art to make and use the claimed subject matter.

This rejection is respectfully traversed.

Relevant law

To satisfy the enablement requirement of 35 U.S.C § 112, first paragraph, the specification must teach one of skill in the art to make and use the invention without undue experimentation. Atlas Powder Co. v. E.I. DuPont de Nemours, 750 F.2d 1569, 224 USPQ 409 (1984). This requirement can be met by providing sufficient disclosure, either through illustrative examples or terminology, to teach one of skill in the art how to make and how to use the claimed subject matter without undue experimentation. This clause does not require "a specific example of everything *within the scope* of a broad claim." In re Anderson, 176 USPQ 331, at 333 (CCPA 1973), emphasis in original. Rather, the requirements of § 112, first paragraph "can be fulfilled by the use of illustrative examples **or** by broad terminology." In re Marzocci et al., 469 USPQ 367 (CCPA 1971)(emphasis added).

Further, because "it is manifestly impracticable for an applicant who discloses a generic invention to give an example of every species falling within it, or even to name every such species, it is sufficient if the disclosure teaches those skilled in the art what the invention is and how to practice it." In re Grimme, Keil and Schmitz, 124 USPQ 449, 502 (CCPA 1960). Thus, there is no doubt that a patentee's invention may be broader than the particular embodiment shown in the specification. A patentee not only is entitled to

narrow claims particularly directed to the preferred embodiment, but also to broad claims that define the invention without a reference to specific instrumentalities. Smith v. Snow, 294 U.S. 1, 11, 24 USPQ 26, 30 (1935).

Thus, there is no requirement for disclosure of every species within a genus. Applicant is entitled to claims are commensurate in scope not only with what applicant has specifically exemplified, but commensurate in scope with that which one of skill in the art could obtain by virtue of that which the applicant has disclosed.

The inquiry with respect to scope of enablement under 35 U.S.C. §112, first paragraph, is whether it would require undue experimentation to make and use the claimed invention. A considerable amount of experimentation is permissible, particularly if it is routine experimentation. The amount of experimentation that is permissible depends upon a number of factors, which include: the quantity of experimentation necessary, the amount of direction or guidance presented, the presence or absence of working examples, the nature of the invention, the state of the prior art, the relative skill of those in the art, the predictability of the art, and the breadth of t he claims. Ex parte Forman, 230 USPQ 546 (Bd. Pat. App. & Int'f 1986); see also In re Wands, 8 USPQ2d 1400 (Fed. Cir. 1988).

Analysis

As demonstrated below, it would not require undue experimentation to select the variant domains of the fusion proteins that are within the scope of the instant claims, given what is known to those of skill in the art and is taught by the the specification regarding the structural and functional characterization of each of domains. Furthermore, it would not require undue experimentation to prepare the fusion proteins as claimed.

As demonstrated below, the teachings of the specification, when taken in conjunction with what is known to one of skill in the art, are such that it would require no undue experimentation to (i) select variant domains having the structural and functional characteristics of ligand binding domains, zinc finger

peptide DNA binding domains, or transcription regulation domains domain that recognizes a specific growth factor or cytokine receptor expressed on a target cell; (ii) isolate and screen the variant domains for their structural and functional limitations as claimed; (iii) construct the fusion proteins as claimed; and (iv) screen for the DNA binding specificity and transcription regulation activity of the constructed fusion proteins.

The level of knowledge and skill in the construction, expression and assay of the claimed fusion proteins was so high as of the effective filing date that it would not have required extensive experimentation by one of skill in the art to substitute a domain variant provided by the methods in the working examples and the publications incorporated herein by reference for the exemplified fusion proteins of the instant application, nor would it have required extensive experimentation to express or assay the fusion proteins where such substitutions have been made.

Scope of the claims

The claims are directed to fusion proteins that function as ligand activated transcriptional regulators. These fusion proteins contain a nucleotide binding domain operatively linked to a ligand binding domain from an intracellular receptor. Depednent claims specify that the fusion proteins further comprise an operatively linked transcription regulating domain. The nucleotide binding domain set forth in the claims is a polydactyl zinc-finger peptide or modular portion thereof that is assembled from one or more modular units wherein each modular unit specifically interacts with a contiguous nucleotide sequence of at least about 3 nucleotides as is known for zinc-finger peptides and other such nucleotide binding domains. Dependent clais recite that the ligand binding domain portion can be from a nuclear hormone receptor; other claims recite that it can be modified to change its selectivity compared to the native hormone receptor. Dependent claims further specify that in the claimed fusion proteins, the zinc-finger peptide can contain a zinc finger or a variant thereof that specifically binds to a targeted nucleic acid molecule, that the hormone receptor

can be a progesterone or estrogen receptor variant that has selectivity and sensitivity for endogenous and exogenous ligands that differ from its native ligands, and the transcription regulating domain can be a variant, derivative, multimer or other combination of known transcription regulating domains.

Teachings of the specification

The classes of molecules belonging to each of the domains of the fusion proteins claimed in the instant application are well characterized in the specification and are known to those of skill in the art, as are the modifications that render variants that retain the structural and functional characteristics of each of these domains. Contrary to the Examiner's assertion, the specification provides the common structural and functional features that characterize each domain, and the specification further provides specific guidance, with numerous examples, as to the preparation of variants that possess the requisite structural and functional characteristics of the instantly claimed fusion proteins.

The specification provides details of the types of modifications that constitute variants that are within the scope of the claims, as well as means by which to identify such variants. While the Examiner's suggestion that single amino acid substitutions may result in dramatic structural and functional changes in a protein is not disputed, the instant specification describes in great detail the selection of variants having the structural and functional limitations of the generic claim, and such selection of variants of each domain as provided in the specification is known to those of skill in the art. Furthermore, the specification provides guidance for selection of domains and variants that have the requisite properties.

As discussed, the specification provides the types of modifications that would constitute variants that are within the scope of the claims, as well as means by which to identify such variants. The specification describes that variants of the ligand binding domain have altered selectivity and sensitivity for endogenous or exogenous ligands, such as a therapeutic ligand, and the specification teaches how ligand binding domains, well characterized and known

to those of skill in the art, may be modified to alter ligand sensitivity. The specification also teaches the types of sequence variants that retain the properties of a zinc finger binding domain or a transcription regulation domain, and the selection of these variants.

For example, page 32, line 7 to page 33, line 11 of the specification provides methods, known to those of skill in the art, to prepare and characterize variants of the ligand binding domain, including specific changes that will provide altered endogenous or exogenous ligand specificity as desired. Page 33, line 12 to page 50, line 2 of the specification provides in exquisite detail and incorporates by reference what was known to those of skill in the art at the time of filing of the application concerning zinc finger proteins, the modular nature of zinc finger proteins wherein each zinc finger specifically recognizes a 3 nucleotide sequence, the types of zinc finger proteins, specific changes that provide variant zinc finger peptides that retain the characteristics of recognizing zinc finger DNA binding motifs, how to construct, isolate or synthesize such variants, and how to screen for such variants. Page 50, line 3 to page 52, line 5 of the specification discloses known transcriptional regulatory domains and selection and modifications thereof. At, for example, page 31, line 6, to page 32, line 6, the specification teaches how to construct the claimed fusion proteins from the various domains and their variants. In addition, numerous working examples, discussed below, are provided throughout the specification, as are exemplary fusion proteins, encoded by SEQ ID NOS. 1-18.

The specification states (pages 37-38), for example:

> DNA recognition in each of the three zinc finger domains of this protein is mediated by residues in the N-terminus of the α-helix contacting primarily three nucleotides on a single strand of the DNA. The operator binding site for this three finger protein is 5'-GCGTGGGCG-'3 (finger-2 subsite is underlined). *Structural studies of Zif268 and other related zinc finger-DNA complexes have shown that residues from primarily three positions on the α-helix, -1, 3, and 6, are involved in specific base contacts. Typically, the residue at position -1 of the α-helix contacts the 3' base of that finger's subsite while positions 3 and 6 contact the middle base and the 5' base, respectively*. [emphasis added]

Thus, the fact that three nucleotides are required was known to those of skill in as well as the requisites for zinc finger binding thereto.

Furthermore, as discussed above methods for preparation of synthetic zinc fingers (as well as other DBD) are described in the application (see, *e.g.*, page 38 *et seq.*):

> Synthetic zinc fingers can be assembled based upon known sequence specificities. A large number of zinc finger-nucleotide binding polypeptides were made and tested for binding specificity against target nucleotides containing a GNN triplet. The data show that a striking conservation of all three of the primary DNA contact positions (-1, 3, and 6) was observed for virtually all the clones of a given target (see, Example 1, see, also U.S. application Serial No. 09/173,941, filed 16 October, 1998, published as International PCT application No. WO 00/23464).
>
> In order to select a family of zinc finger domains recognizing the 5'-GNN-3' subset of sequences, two highly diverse zinc finger libraries were constructed in the phage display vector pComb3H (Barbas *et al.* (1991) *Proc. Natl. Acad. Sci. USA 88*:7978-7982; Rader *et al.* (1997) *Curr. Opin. Biotechnol. 8*:503-508). Both libraries involved randomization of residues within the α-helix of finger 2 of C7, a variant of Zif268 (Wu *et al.* (1995) *Proc. Natl. Acad. Sci. U.S.A. 92*:344-348). Library 1 was constructed by randomization of positions -1,1,2,3,5,6 using a NNK doping strategy while library 2 was constructed using a VNS doping strategy with randomization of positions -2,-1,1,2,3,5,6. The NNK doping strategy allows for all amino acid combinations within 32 codons while VNS precludes Tyr, Phe, Cys and all stop codons in its 24 codon set. The libraries contained 4.4 x 10^9 and 3.5 x 10^9 members, respectively, each capable of recognizing sequences of the 5'-GCGNNNGCG-3' type. The size of the NNK library ensured that it could be surveyed with 99% confidence while the VNS library was highly diverse but somewhat incomplete. These libraries are, however, significantly larger than previously reported zinc finger libraries (International PCT application No. WO 09/54311; Choo *et al.* (1994) *Proc Natl Acad Sci U S A 91*:11163-7; Greisman *et al.* (1997) *Science 275*:657-661; Rebar *et al.* (1994) *Science 263*:671-673; Jamieson *et al.* (1994) *Biochemistry 33*:5689-5695; Jamieson *et al.* 1996) *Proc. Natl. Acad. Sci. U.S.A. 93*:12834-12839; Isalan *et al.* (1998) *Biochemistry 37*:12026-12033; and U.S. Patent No. 5,789,538). Seven rounds of selection were performed on the zinc finger displaying-phage with each of the 16 5'-GCGGNNGCG-3' biotinylated hairpin DNAs targets using a solution binding protocol. Stringency was increased in each round by the addition of competitor DNA. Sheared herring sperm DNA was provided

for selection against phage that bound non-specifically to DNA. Stringent selective pressure for sequence specificity was obtained by providing DNAs of the 5'-GCGNNNGCG-3' types as specific competitors. Excess DNA of the 5'-GCGGNNGCG-3' type was added to provide even more stringent selection against binding to DNAs with single or double base changes as compared to the biotinylated target. Phage binding to the single biotinylated DNA target sequence were recovered using streptavidin coated beads. In some cases the selection process was repeated. The data show that these domains are functionally modular and can be recombined with one another to create proteins capable of binding to 18-bp sequences with subnanomolar affinity. The resulting family of zinc finger domains described herein is sufficient for the construction of 17 million proteins that bind to the 5'-$(GNN)_6$-3' family of DNA sequences.

Also impressive amino acid conservation was been observed for recognition of the same nucleotide in different targets. For example, Asn in position 3 (Asn3) virtually always selects to recognize adenine in the middle position, whether in the context of GAG, GAA, GAT, or GAC. Gln-1 and Arg-1 were always selected to recognize adenine or guanine, respectively, in the 3' position regardless of context. Amide side chain based recognition of adenine by Gln or Asn is well documented in structural studies as is the Arg guanidinium side chain to guanine contact with a 3' or 5' guanine (see, *e.g.*, Elrod-Erickson *et al.* (1998) *Structure 6*:451-464).

More often, however, two or three amino acids are selected for nucleotide recognition. His3 or Lys3 (and to a lesser extent, Gly3) are selected for the recognition of a middle guanine. Ser3 and Ala3 are selected to recognize a middle thymine. Thr3, Asp3, and Glu3 are selected to recognize a middle cytosine. Asp and Glu were are selected in position -1 to recognize a 3' cytosine, while Thr-1 and Ser-1 are selected to recognize a 3' thymine.

Specific recognition of many nucleotides can best accomplished using motifs, rather than a single amino acid. For example, the best specification of a 3' guanine is achieved using the combination of Arg-1, Ser1, and Asp2 (the RSD motif). By using Val5 and Arg6 to specify a 5' guanine, recognition of subsites GGG, GAG, GTG, and GCG can be accomplished using a common helix structure (SRSD-X-LVR) differing only in the position 3 residue (Lys3 for GGG, Asn3 for GAG, Glu3 for GTG, and Asp3 for GCG). Similarly, 3' thymine is specified using Thr-1, Ser1, and Gly2 in the final clones(the TSG motif). Further, a 3' cytosine can be specified using Asp-1, Pro1, and Gly2 (the DPG motif) except when the

subsite is GCC; Pro1 is not tolerated by this subsite. Specification of a 3' adenine is with Gln-1, Ser1, Ser2 in two clones (QSS motif).

The data (see, Table 1 in Example) show that all possible GNN triplet sequences can be recognized with exquisite specificity by zinc finger domains. Optimized zinc finger domains can discriminate single base differences by greater than 100-fold loss in affinity. While many of the amino acids found in the optimized proteins at the key contact positions -1,3, and 6 are those that are consistent with a simple code of recognition, it has been discovered that optimal specific recognition is sensitive to the context in which these residues are presented. Residues at positions 1,2, and 5 have been found to be critical for specific recognition.

Further the data demonstrate that sequence motifs at positions -1,1, and 2 rather than the simple identity of the position 1 residue are required for highly specific recognition of the 3' base. These residues likely provide the proper stereo-chemical context for interactions of the helix in terms of recognition of specific bases and in the exclusion of other bases, the net result being highly specific interactions. Ready recombination of the disclosed domains then allows for the creation of proteins, typically polypdactyl proteins, of defined specificity precluding the need to develop phage display libraries in their generation. Such family of zinc finger domains is sufficient for the construction of 16 or 17 million proteins that bind to the 5'-$(GNN)_6$-3' family of DNA sequences.

Hence the specification establishes that those of skill in the art have done an extensive amount of experimentation and preparatio of synthetic zinc fingers and describes in great detail how to effect such modifications.

The specification further describes (see, *e.g.*, page 41 *et seq.*) modifciations of zinc fingers and describes methods for screening for variant zinc fingers (page 45 *et seq.*):

The zinc finger-nucleotide binding peptide domain can be derived or produced from a wild type zinc finger protein by truncation or expansion, or as a variant of the wild type-derived polypeptide by a process of site directed mutagenesis, or by a combination of the procedures (see, *e.g.*, U.S. Patent No. 5,789.538, which describes methods for design and construction of zinc finger peptides). Mutagenesis can be performed to replace non-conserved residues in one or more of the repeats of the consensus sequence. Truncated zinc finger-nucleotide binding proteins can also be mutagenized.

DNA encoding the zinc finger-nucleotide binding proteins, including native, truncated, and expanded polypeptides, can be obtained by several methods. For example, the DNA can be isolated using hybridization procedures which are well known in the art. These include, but are not limited to: (1) hybridization of probes to genomic or cDNA libraries to detect shared nucleotide sequences; (2) antibody screening of expression libraries to detect shared structural features; and (3) synthesis by the polymerase chain reaction (PCR). RNA can be obtained by methods known in the art (seem *e.g.*, *Current Protocols in Molecular Biology*, 1988, Ed. Ausubel, *et al.*, Greene Publish. Assoc. & Wiley Interscience).

DNA encoding zinc finger-nucleotide binding proteins also can be obtained by: (1) isolation of a double-stranded DNA sequence from the genomic DNA; (2) chemical manufacture of a DNA sequence to provide the necessary codons for the polypeptide of interest; and (3) *in vitro* synthesis of a double-stranded DNA sequence by reverse transcription of mRNA isolated from a eukaryotic donor cell. In the latter case, a double-stranded DNA complement of mRNA is eventually formed which is generally referred to as cDNA. Of these three methods the isolation of genomic DNA is the least common. This is especially true when it is desirable to obtain the microbial expression of mammalian polypeptides due to the presence of introns.

For obtaining zinc finger derived-DNA binding polypeptides, the synthesis of DNA sequences is frequently the method of choice when the entire sequence of amino acid residues of the desired polypeptide product is known. When the entire sequence of amino acid residues of the desired polypeptide is not known, the direct synthesis of DNA sequences is not possible and the method of choice is the formation of cDNA sequences. Among the standard procedures for isolating cDNA sequences of interest is the formation of plasmid-carrying cDNA libraries which are derived from reverse transcription of mRNA which is abundant in donor cells that have a high level of genetic expression. When used in combination with polymerase chain reaction technology, even rare expression products can be cloned. In those cases where significant portions of the amino acid sequence of the polypeptide are known, the production of labeled single or double-stranded DNA or RNA probe sequences duplicating a sequence putatively present in the target cDNA may be employed in DNA/DNA hybridization procedures which are carried out on cloned copies of the cDNA which have been denatured into a single-stranded form (Jay, *et al.*, *Nucleic Acid Research*, 11:2325, 1983).

Hybridization procedures are useful for the screening of recombinant clones by using labeled mixed synthetic oligonucleotide

probes where each probe is potentially the complete complement of a specific DNA sequence in the hybridization sample which includes a heterogeneous mixture of denatured double-stranded DNA. For such screening, hybridization is preferably performed on either single-stranded DNA or denatured double-stranded DNA. Hybridization is particularly useful in the detection of cDNA clones derived from sources where an extremely low amount of mRNA sequences relating to the polypeptide of interest are present. By using stringent hybridization conditions directed to avoid non-specific binding, it is possible, for example, to allow the autoradiographic visualization of a specific cDNA clone by the hybridization of the target DNA to that single probe in the mixture which is its complete complement (Wallace, *et al., Nucleic Acid Research*, 9:879, 1981; Maniatis, *et al., Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory, 1982).

Screening procedures that rely on nucleic acid hybridization make it possible to isolate any gene sequence from any organism, provided the appropriate probe is available. Oligonucleotide probes, which correspond to a part of the sequence encoding the protein in question, can be synthesized chemically. This requires that short, oligopeptide stretches of amino acid sequence must be known. The DNA sequence encoding the protein can be deduced from the genetic code, however, the degeneracy of the code must be taken into account. It is possible to perform a mixed addition reaction when the sequence is degenerate. This includes a heterogeneous mixture of denatured double-stranded DNA. For such screening, hybridization is preferably performed on either single-stranded DNA or denatured double-stranded DNA.

A cDNA expression library, such as lambda gt11, can be screened indirectly for zinc finger-nucleotide binding protein or for the zinc finger derived polypeptide having at least one epitope, using antibodies specific for the zinc finger-nucleotide binding protein. Such antibodies can be either polyclonally or monoclonally derived and used to detect expression product indicative of the presence of zinc finger-nucleotide binding protein cDNA. Alternatively, binding of the derived polypeptides to DNA targets can be assayed by incorporated radiolabeled DNA into the target site and testing for retardation of electrophoretic mobility as compared with unbound target site.

A preferred vector used for identification of truncated and/or mutagenized zinc finger-nucleotide binding polypeptides is a recombinant DNA molecule containing a nucleotide sequence that codes for and is capable of expressing a fusion polypeptide containing, in the direction of amino- to carboxy-terminus, (1) a prokaryotic secretion signal domain, (2)

a heterologous polypeptide, and (3) a filamentous phage membrane anchor domain. The vector includes DNA expression control sequences for expressing the fusion polypeptide, preferably prokaryotic control sequences.

Since the DNA sequences provided herein encode essentially all or part of an zinc finger-nucleotide binding protein, it is routine to prepare, subclone, and express the truncated polypeptide fragments of DNA from this or corresponding DNA sequences. Alternatively, by using the DNA fragments disclosed herein, which define the zinc finger-nucleotide binding polypeptides, it is possible, in conjunction with known techniques, to determine the DNA sequences encoding the entire zinc finger-nucleotide binding protein. Such techniques are described in U.S. 4,394,443 and U.S. 4,446,235, which are incorporated herein by reference.

In addition to modifications in the amino acids making up the zinc finger, the zinc finger derived polypeptide can contain more or less than the full amount of fingers contained in the wild type protein from which it is derived. Minor modifications of the primary amino acid sequence may result in proteins which have substantially equivalent activity compared to the zinc finger derived-binding protein described herein. Such modifications may be deliberate, as by site-directed mutagenesis, or may be spontaneous. All proteins produced by these modifications are included herein as long as zinc finger-nucleotide binding protein activity exists.

e. Screening of varint zinc finger and other DBD peptides

Any method known to those of skill in the art for identification of functional modular domains derived from zinc fingers and combinations thereof can be employed. An exemplary method for identifying variants of zinc fingers or other polypeptides that bind to zinc finger binding motifs is provided. Components used in the method include a nucleic acid molecule encoding a putative or modified zinc finger peptide operably linked to a first inducible promoter and a reporter gene operably linked to a second inducible promoter and a zinc finger-nucleotide binding motif, wherein the incubating is carried out under conditions sufficient to allow the components to interact, and measuring the affect of the putative DBD peptide on the expression of the reporter gene is provided.

For exampole, a first inducible promoter, such as the arabinose promoter, is operably linked to the nucleotide sequence encoding the putative DBD polypeptide. A second inducible promoter, such as the

lactose promoter, is operably linked to a zinc finger derived-DNA binding motif followed by a reporter gene, such as β-galactosidase. Incubation of the components may be *in vitro* or *in vivo*. *In vivo* incubation may include prokaryotic or eukaryotic systems, such as *E.coli* or COS cells, respectively. Conditions that allow the assay to proceed include incubation in the presence of a substance, such as arabinose and lactose, which activate the first and second inducible promoters, respectively, thereby allowing expression of the nucleotide sequence encoding the putative trans-modulating protein nucleotide sequence. Determination of whether the putative modulating protein binds to the zinc finger-nucleotide binding motif, which is operably linked to the second inducible promoter, and affects its activity is measured by the expression of the reporter gene. For example, if the reporter gene is β-galactosidase, the presence of blue or white plaques indicates whether the putative modulating protein enhances or inhibits, respectively, gene expression from the promoter. Other commonly used assays to assess the function from a promoter, including chloramphenicol acetyl transferase (CAT) assay, are known to those of skill in the art. Prokaryote and eukaryote systems can be used.

As discussed above, Example 1 provides an illustration of modification of Zif268 as described above. Therefore, in another embodiment, a ligand activated transcriptional regulator polypeptide variant containing at least two zinc finger modules that bind to an HIV sequence and modulates the function of the HIV sequence, for example, the HIV promoter sequence is provided.

In another embodiment, zinc finger proteins can be manipulated to recognize and bind to extended target sequences. For example, zinc finger proteins containing from about 2 to 20 zinc fingers Zif(2) to Zif(20), and preferably from about 2 to 12 zinc fingers, may be fused to the leucine zipper domains of the Jun/Fos proteins, prototypical members of the bZIP family of proteins (O'Shea *et al.* (1991) *Science 254*:539). Alternatively, zinc finger proteins can be fused to other proteins which are capable of forming heterodimers and contain dimerization domains. Such proteins are known to those of skill in the art.

The Jun/Fos leucine zippers are described for illustrative purposes and preferentially form heterodimers and allow for the recognition of 12 to 72 base pairs. Henceforth, Jun/Fos refer to the leucine zipper domains of these proteins. Zinc finger proteins are fused to Jun, and independently to Fos by methods commonly used in the art to link proteins. Following purification, the Zif-Jun and Zif-Fos constructs, the proteins are mixed to spontaneously form a Zif-Jun/Zif-Fos heterodimer.

Alternatively, coexpression of the genes encoding these proteins results in the formation of Zif-Jun/Zif-Fos heterodimers *in vivo*. Fusion of the heterodimer with an N-terminal nuclear localization signal allows for targeting of expression to the nucleus (Calderon, *et al, Cell*, 41:499, 1982). Activation domains may also be incorporated into one or each of the leucine zipper fusion constructs to produce activators of transcription (Sadowski *et al.* (1992) *Gene 118*:137). These dimeric constructs then allow for specific activation or repression of transcription. These heterodimeric Zif constructs are advantageous since they allow for recognition of palindromic sequences (if the fingers on Jun and Fos recognize the same DNA/RNA sequence) or extended asymmetric sequences (if the fingers on Jun and Fos recognize different DNA/RNA sequences). For example the palindromic sequence

5′ - GGC CCA CGC $\{N\}_x$ GCG TGG GCG - 3′

3′ - GCG GGT GCG $\{N\}_x$ CGC ACC CGC - 5′ (SEQ ID NO: 20)

is recognized by the Zif268-Fos/Zif268 Jun dimer (x is any number). The spacing between subsites is determined by the site of fusion of Zif with the Jun or Fos zipper domains and the length of the linker between the Zif and zipper domains. Subsite spacing is determined by a binding site selection method as is common to those skilled in the art (Thiesen *et al.* (1990) *Nucleic Acids Research,* 18:3203, 1990). Example of the recognition of an extended asymmetric sequence is shown by the Zif$(C7)_6$-Jun/Zif-268-Fos dimer. This protein includes 6 fingers of the C7 type (EXAMPLE 11) linked to Jun and three fingers of Zif268 linked to Fos, and recognizes the extended sequence:

5′ - CGC CGC CGC CGC CGC CGC $\{N\}_x$ GCG TGG GCG - 3′

3′ - GCG GCG GCG GCG GCG GCG $\{N\}_x$ CGC ACC CGC - 5′

(SEQ ID NO: 21)

In another embodiment, attachment of chelating groups to Zif proteins is preferably facilitated by the incorporation of a Cysteine (Cys) residue between the initial Methionine (Met) and the first Tyrosine (Tyr) of the protein. The Cys is then alkylated with chelators known to those skilled in the art, for example, EDTA derivatives as described (Sigman (1990) *Biochemistry*, 29:9097). Alternatively the sequence Gly-Gly-His can be made as the most amino terminal residues since an amino terminus composed of the residues has been described to chelate Cu^{+2} (Mack *et al.* (1988) *J. Am. Chem. Soc. 110*:7572). Preferred metal ions include Cu^{+2}, Ce^{+3} (Takasaki and Chin (1994) *J. Am. Chem. Soc. 116*:1121, 1994) Zn^{+2}, Cd^{+2}, Pb^{+2}, Fe^{+2} (Schnaith *et al.* (1994) *Proc. Natl. Acad. Sci., USA 91*:569, 1994), Fe^{+3}, Ni^{+2}, Ni^{+3}, La^{+3}, Eu^{+3} (Hall*et al.* (1994) *Chemistry and Biology 1*:185), Gd^{+3}, Tb^{+3}, Lu^{+3}, Mn^{+2}, Mg^{+2}. Cleavage with chelated metals is generally performed in the presence of oxidizing agents such as O_2, hydrogen peroxide H_2O_2 and reducing agents such as thiols and ascorbate. The site and strand (+ or - site) of

cleavage is determined empirically (Mack *et al.* (1988) *J. Am. Chem. Soc 110*:7572, 1988) and is dependent on the position of the Cys between the Met and the Tyr preceding the first finger. In the protein Met (AA) Tyr-$(Zif)_{1-12}$, the chelate becomes Met-$(AA)_{x1}$ Cys-Chelate-$(AA)_{x2}$-Tyr-$(Zif)_{1-12}$, where AA = any amino acid and x = the number of amino acids. Dimeric zif constructs of the type Zif-Jun/Zif-Fos are preferred for cleavage at two sites within the target oligonucleotide or at a single long target site. In the case where double stranded cleavage is desired, Jun and Fos containing proteins are labelled with chelators and cleavage is performed by methods known to those skilled in the art. In this case, a staggered double-stranded cut analogous to that produced by restriction enzymes is generated.

Following mutagenesis and selection of variants of the Zif268 protein in which the finger 1 specificity or affinity is modified, proteins carrying multiple copies of the finger may be constructed using the TGEKP linker sequence by methods known in the art. For example, the C7 finger may be constructed according to the scheme: MKLLEPYACPVESCDRRFSKSADLKRHIRHTGEKP- (SEQ ID NO: 22) $(YACPVESCDRRFSKSADLKHIRIHTGEKP)_{1-11}$, (SEQ ID NO: 23) where the sequence of the last linker is subject to change since it is at the terminus and not involved in linking two fingers together. This protein binds the designed target sequence GCG-GCG-GCG in the oligonucleotide hairpin CCT-CGC-CGC-CGC-GGG-TTT-TCC-CGC-GCC-CCC GAG G (SEQ ID NO: 24) with an affinity of 9nM, as compared to an affinity of 300 nM for an oligonucleotide encoding the GCG-TGG-GCG sequence (as determined by surface plasmon resonance studies). Fingers used need not be identical and may be mixed and matched to produce proteins which recognize a desired target sequence. These may also be used with leucine zippers (*e.g.,* Fos/Jun) or other heterodimers to produce proteins with extended sequence recognition.

In addition to producing polymers of finger 1, the entire three finger Zif268 and modified versions therein may be fused using the consensus linker TGEKP to produce proteins with extended recognition sites. For example, the protein Zif268-Zif268 can be produced in which the natural protein has been fused to itself using the TGEKP linker. This protein now binds the sequence GCG-TGG-GCG-GCG-TGG-GCG. Therefore modifications within the three fingers of Zif268 or other zinc finger proteins known in the art may be fused together to form a protein which recognizes extended sequences. These new zinc proteins may also be used in combination with leucine zippers if desired.

The specification provides similar guidance for each domain in the claimed fusion proteins.

Level of skill

The level of skill in this art is recognized to be high (see, *e.g.*, Ex parte Forman, 230 USPQ 546 (Bd. Pat. App. & Int'f 1986)). The numerous articles and patents made of record in this application address a highly skilled audience and further evidence the high level of skill in this art.

Knowledge of those of skill in the art

At the time of the effective filing date of this application and before, the skilled artisan knew the biochemical and structural characteristics and properties of receptors and proteins forming the domains of the instantly claimed fusion proteins. Further, there was a large body of literature, set forth below and incorporated in the instant specification by reference, that was directed to the identity, structure and function of each of the individual domains of the claimed fusion proteins. Moreover, the recognition domains and/or ligands through which each of the domains elicit their effects were well known. Also known to those of skill in the art were methods of selecting, modifying and screening for each of the domain variants to obtain the structural and functional properties of ligand binding, target sequence recognition and binding, and transcription regulation that are desired.

For instance, means to modify and test the specificity of hormone receptor ligand binding domains and to identify ligands therefor were known (see, U.S. Patent No. 5,874,534; U.S. Patent No. 5,935,934; and International PCT application No. 98/18925, which is based on U.S. provisional application Serial No. 60/029,964; International PCT application No. 96/40911, which is based on U.S. application Serial No. 08/479,913). Exemplary ligand binding domain modifications that lead to desired properties, such as preferential interaction with non-natural ligands, were also known to those of skill in the art (see, *e.g.*, U.S. Patent No. 5,874,534; U.S. Patent No. 5,935,934; U.S. Patent No. 5,364,791; and International PCT application No. 98/18925, which is based

on U.S. provisional application Serial No. 60/029,964; International PCT application No. 96/40911, which is based on U.S. application Serial No. 08/479,913) and references cited therein.

For example, methods and rules for preparation and selection of zinc fingers based upon the C2H2 class of zinc fingers with unique specificity were known (see, *e.g.*, International PCT application No. WO 98/54311 and International PCT application No. 95/19431; see, also U.S. Patent No. 5,789,538; Beerli *et al.* (1999) *Proc. Natl. Acad. Sci. U.S.A. 96*:2758-2763; Beerli *et al.* (1995) *Proc. Natl. Acad. Sci. U.S.A. 95*:14628-14633; see, also U.S. application Serial No. 09/173,941, filed 16 October, 1998, published as International PCT application No. WO 00/23464). It was also known that zinc finger variants can be prepared by identifying a zinc finger or modular unit thereof, creating an expression library, such as a phage display library (see, *e.g.*, International PCT application No. WO 98/54311, Barbas *et al.* (1991) *Methods 2*:119; Barbas *et al.* (1992) *Proc. Natl. Acad. Sci. U.S.A. 89*:4457), encoding polypeptide variants of the zinc finger or modular unit thereof, expressing the library in a host and screening for variant peptides having a desired specificity.

It was known to those of skill in the art that a zinc finger-nucleotide binding peptide domain contains a unique heptamer (contiguous sequence of 7 amino acid residues) within the α-helical domain of the polypeptide, which heptameric sequence determines binding specificity to a target nucleotide. It was also known that three zinc finger domains can bind 9 bp of contiguous DNA sequence (Pavletich *et al.* (1991) *Science 252*:809-817; Swirnoff *et al.* (1995) *Mol. Cell. Biol. 15*:2275-2287). Furthermore, it was known that while recognition of 9 bp of sequence is insufficient to specify a unique site in a complex genome, proteins containing six zinc finger domains can specify 18-bp recognition (Liu *et al.* (1997) *Proc. Natl. Acad. Sci. USA 94*:5525-5530). An 18-bp address made up of modular units was known to be of sufficient complexity to specify a single site within all known genomes (see, published International PCT application No. WO 98/54311). Rules for constructing Zinc

finger arrays that bind to a particular DNA sequence were known (see, *e.g.*, International PCT application No. WO 98/54311, which is based on U.S. application Serial No. 08/863,813; International PCT application No. 95/19431, which is based on U.S. application Serial Nos. 08/183,119 and 08/312,604).

The murine Cys_2-His_2 zinc finger protein Zif268 had been used for construction of phage display libraries (Wu *et al.* (1995) *Proc. Natl. Acad. Sci. U.S.A. 92*:344-348). Zif268 was a structurally well characterized zinc-finger proteins (Pavletich, *et al.* (1991) *Science 252*:809-817; Elrod-Erickson *et al.* (1996) *Structure 4*:1171-1180; Swirnoff *et al.* (1995) *Mol. Cell. Biol. 15*:2275-2287). DNA recognition in each of the three zinc finger domains of this protein was known to be mediated by residues in the N-terminus of the α-helix contacting primarily three nucleotides on a single strand of the DNA. The operator binding site for this three finger protein was identified as 5'-GCGTGGGCG-'3 (finger-2 subsite is underlined). Structural studies of Zif268 and other related zinc finger-DNA complexes had shown that residues from primarily three positions on the α-helix, -1, 3, and 6, are involved in specific base contacts.

Also known to those of skill in the art at the time of filing of the instant application was the construction of highly diverse zinc finger libraries in the phage display vector pComb3H (Barbas *et al.* (1991) *Proc. Natl. Acad. Sci. USA 88*:7978-7982; Rader *et al.* (1997) *Curr. Opin. Biotechnol. 8*:503-508). Both libraries involved randomization of residues within the α-helix of finger 2 of C7, a variant of Zif268 (Wu *et al.* (1995) *Proc. Natl. Acad. Sci. U.S.A. 92*:344-348). Techniques for modifying zinc finger peptides were known (see, *e.g.*, U.S. Patent No. 5,789,538), as were binding site selection methods (Thiesen *et al.* (1990) *Nucleic Acids Research,* 18:3203, 1990).

Transcriptional regulatory domains were also known to those of skill in the art, as were transcriptional repressors and their structural and functional characteristics (Sgouras *et al.* (1995) *EMBO J. 14*:4781-4793, ERD domain mediating the antagonistic effect of ERF on the activity of transcription factors

of the *ets* family; Margolin *et al.* (1994) *Proc. Natl. Acad. Sci. USA 91*:4509-4513; Pengue *et al.* (1996) *Proc. Natl. Acad. Sci. USA 93*:1015-1020; Friedman *et al.* (1996) *Genes & Dev. 10*:2067-2078, KRAB repressor domain; Ayer *et al.* (1996) *Mol. Cell. Biol. 16*:5772-5781, repression by histone deacetylation).

Presence of working examples

The specification provides numerous working examples and descriptions of the construction and expression of the claimed fusion proteins. Example 1 beginning at page 67 of the specification describes in great detail the construction and testing of exemplary zinc finger domains. Example 1 shows how variants that specifically recognize the "GNN" target triplet nucleotide sequence of a zinc finger modular unit can be selected. Example 1 also provides numerous exemplary target sequences from various regions of the *erbB-2* and *integrin β3* promoters that provide zinc finger recognition motifs. These target sequences (Table 2 of Example 1) along with the exemplary zinc finger modular units and their recognition sequences provided in Example 1 (Table 1 of Example 1) can be extrapolated to similarly tailor zinc finger binding domains to interact with selected sequences of targets other than the *erbB-2* and *integrin β3* promoter regions that are exemplified herein.

Example 2 beginning at page 84 provides the construction of fusion proteins containing zinc finger domains and transcriptional repressors and activators. Example 2 also shows regulation of *erbB-2* and *integrin β3* promoter activity using these exemplary constructs. Examples 3-6 beginning at page 87 of the specification and Examples 9-11 beginning at page 105 of the specification provide the construction of fusion proteins containing progesterone or estrogen receptor ligand binding domains either fused to zinc finger DNA binding domains alone or to zinc finger domains and transcriptional activators or repressors. The fusion proteins provided in the aforementioned examples were tested for their ability to regulate *erbB-2* (Examples 9 and 11), *integrin β3* (Examples 9 and 11), SV40 (Examples 3, 5 and 10) and CMV (Example 6)

promoters and various palindromic oligonucleotide target sequences (Example 4).

Example 5 at page 97 of the specification provides the ligand dependent regulation of transgene expression by exemplary fusion proteins, and Examples 6-8 and 16-18 beginning at page 100 and page 119, respectively, of the specification provides exemplary structural characterizations and evaluations of the correlating regulatory activity of the individual domains and the fusion protein constructs. Example 14 at page 113 of the specification demonstrates how the estrogen receptor ligand binding domain may be modified to obtain variants with altered ligand selectivity.

Predictability

Since the individual components and functioning of the fusion proteins are known and methods and assays for modifying receptor binding domains are known, it is not unpredictable that variants can be prepared. The Examiner cites Voet *et al.* (*Biochemistry*, John Wiley & Sons, (1990)), for the proposition that even single amino acid changes can have dramatic effects on a protein's function, structure or architecture. The Examiner concludes that while the claims are enabled for polydactyl zinc finger fusion proteins that bind the erbB2 promoter, they are not enabled for polydactyl zinc finger fusion proteins that bind to any other contiguous nucleotide sequences of at least 3 nucleotides, or a variant of such a fusion protein.

The Voet *et al.* reference, published in 1990, and any broad sweeping statemetn attributed thereto is of no relevance to the enablement claims in an application filed in October, 1999. As discussed above, the specification describes in detail the requisites for each domain, how to identify each domain, and how to make and test variants. The specification demonstrates that much was known about the requisites for binding of DBDs. For example, as discussed above, the specification states (pages 37-38):

> DNA recognition in each of the three zinc finger domains of this protein is mediated by residues in the N-terminus of the α-helix contacting primarily

> three nucleotides on a single strand of the DNA. The operator binding site for this three finger protein is 5'-GCGTGGGCG-'3 (finger-2 subsite is underlined). *Structural studies of Zif268 and other related zinc finger-DNA complexes have shown that residues from primarily three positions on the α-helix, -1, 3, and 6, are involved in specific base contacts. Typically, the residue at position -1 of the α-helix contacts the 3' base of that finger's subsite while positions 3 and 6 contact the middle base and the 5' base, respectively*. [emphasis added]

Thus, the fact that three nucleotides are required was known to those of skill in as well as the requisites for zinc finger binding thereto.

As discussed above, the requisites for binding and for modification of each domain are described in great detail in the specification and modified DBDs are known. Hence, there is no issue of a lack of predictability for preparation of fusion proteins having the requisite properties.

Conclusion

Therefore, in light of the extensive teachings and examples in the specification, the high level of skill of those in this art, the significant knowledge of those of skill in the art, and the breadth of the claims, it would not require undue experimentation for the skilled artisan to make and use the claimed fusion proteins.

Also, since the structural and functional characteristics of the various domain components of the instantly claimed fusion proteins are known, and since their effects can be determined by the standard methods extensively elucidated in the specification, it would be unfair and unduly limiting to require Applicant to limit these claims to a few exemplary sequences. To do so is contrary to the public policy upon which the U.S. patent laws are based. If Applicant is required to limit the claims to only the exemplified fusion proteins, then those of skill in the art could by virtue of the teachings of this application readily practice what is claimed by substituting other ligand binding domains, zinc finger peptides or modular units thereof, and transcription regulation domains, but avoid infringing such limited claims. To permit that is simply not fair. The instant application exemplifies the means for isolation, modification

and screening of domains having the structural and functional limitations as generically claimed, as well as the means for construction and expression of the fusion proteins, and *in vitro* and *in vivo* assays for their sequence specificity and regulatory activity. Having done so, it is now routine to for others to insert other ligand binding, zinc finger, and transcription regulatory domains into the exemplified fusion proteins. Those of skill in the art should not be permitted to make such minor modifications by substitution of a different host and avoid infringing such claims.

REBUTTAL TO SPECIFIC ISSUES RAISED BY THE EXAMINER

1) The specification provides ample guidance as to which of the variant fusion polypeptides provided in the instant application will retain the characteristics of a transcriptional regulator.

The Examiner alleges that the specification provides insufficient guidance as to which of the "myriad of variant fusion polypeptides" will retain the characteristics of a transcriptional regulator. To support this assertion, the Examiner cites Voet *et al.* (*Biochemistry*, John Wiley & Sons, (1990)), which allegedly teaches that a single Glu to Val substitution in the beta subunit of hemoglobin causes the hemoglobin molecules to associate with one another in such a manner that, in homozygous individuals, erythrocytes are altered from their normal discoid shape to a sicke shape characteristic of sickle-cell anemia.

The instant specification provides variant fusion proteins that are not random variations or mutations. Rather, they are selected to retain the structural and functional properties of a ligand activated transcriptional regulator. As noted, Voet *et al.*, published in 1990, is not relevant to the instant claims, since it is not determinative of enablement at the time of filing of the instant application in October, 1999. Furthermore, Voet *et al.* makes no statements regarding the effect of variation in structure on the fusion proteins provided herein.

As discussed above, it would not require undue experimentation to select the variant domains of the fusion proteins that are within the scope of the

instant claims, given what is known to those of skill in the art and taught and exmplified the specification regarding the structural and functional characterization of each of domains. Furthermore, it would not require undue experimentation to prepare the fusion proteins as claimed.

As also discussed above, the specification provides in great detail the types of modifications that constitute variants that are within the scope of the claims, as well as means by which to identify such variants. The specification also provides the types of sequence variants that retain the properties of a zinc finger binding domain or a transcription regulation domain, and the selection of these variants. For example, page 32, line 7 to page 33, line 11 of the specification provides methods, known to those of skill in the art, to prepare and characterize variants of the ligand binding domain, including specific changes that will provide altered endogenous or exogenous ligand specificity as desired. Page 33, line 12 to page 50, line 2 of the specification provides in exquisite detail and incorporates by reference what was known to those of skill in the art at the time of filing of the application concerning zinc finger proteins, the modular nature of zinc finger proteins wherein each zinc finger specifically recognizes a 3 nucleotide sequence, the types of zinc finger proteins, specific changes that provide variant zinc finger peptides that retain the characteristics of recognizing zinc finger DNA binding motifs, how to construct, isolate or synthesize such variants, and how to screen for such variants. Page 50, line 3 to page 52, line 5 of the specification discloses known transcriptional regulatory domains and selection and modifications thereof. At, for example, page 31, line 6, to page 32, line 6, the specification teaches how to construct the claimed fusion proteins from the various domains and their variants. In addition, numerous working examples, discussed above, are provided throughout the specification, as are exemplary fusion proteins, encoded by SEQ ID NOS. 1-18.

Moreover, at the time of the effective filing date of this application and before, there was a large body of literature, set forth below and incorporated in the instant specification by reference, that was directed to the identity,

structure and function of each of the domains of the claimed fusion proteins. Also known to those of skill in the art were methods of selecting, modifying and screening for each of the domain variants to obtain the structural and functional properties of ligand binding, target sequence recognition and binding, and transcription regulation that are desired.

For instance, means to modify and test the specificity of hormone receptor ligand binding domains and to identify ligands therefor were known (see, U.S. Patent No. 5,874,534; U.S. Patent No. 5,935,934; and International PCT application No. 98/18925, which is based on U.S. provisional application Serial No. 60/029,964; International PCT application No. 96/40911, which is based on U.S. application Serial No. 08/479,913). Exemplary ligand binding domain modifications that lead to desired properties, such as preferential interaction with non-natural ligands, were also known to those of skill in the art (see, *e.g.*, U.S. Patent No. 5,874,534; U.S. Patent No. 5,935,934; U.S. Patent No. 5,364,791; and International PCT application No. 98/18925, which is based on U.S. provisional application Serial No. 60/029,964; International PCT application No. 96/40911, which is based on U.S. application Serial No. 08/479,913) and references cited therein. Further, methods and rules for preparation and selection of zinc fingers based upon the C2H2 class of zinc fingers with unique specificity were known (see, *e.g.*, International PCT application No. WO 98/54311 and International PCT application No. 95/19431; see, also U.S. Patent No. 5,789,538; Beerli *et al.* (1999) *Proc. Natl. Acad. Sci. U.S.A. 96*:2758-2763; Beerli *et al.* (1995) *Proc. Natl. Acad. Sci. U.S.A. 95*:14628-14633; see, also U.S. application Serial No. 09/173,941, filed 16 October, 1998, published as International PCT application No. WO 00/23464).

Transcriptional regulatory domains were also known to those of skill in the art, as were transcriptional repressors and their structural and functional characteristics (Sgouras *et al.* (1995) *EMBO J. 14*:4781-4793, ERD domain mediating the antagonistic effect of ERF on the activity of transcription factors of the *ets* family; Margolin *et al.* (1994) *Proc. Natl. Acad. Sci. USA 91*:4509-

4513; Pengue *et al.* (1996) *Proc. Natl. Acad. Sci. USA 93*:1015-1020; Friedman *et al.* (1996) *Genes & Dev. 10*:2067-2078, KRAB repressor domain; Ayer *et al.* (1996) *Mol. Cell. Biol. 16*:5772-5781, repression by histone deacetylation).

Thus there is abundant guidance in the instant specification for the selection of fusion protein variants that possess the structural and functional characteristics of each of the component domains and that are ligand activated transcriptional regulators.

2) The claims are enabled not only for polydactyl zinc finger fusion proteins that bind the *erbB-2* promoter, but also for polydactyl zinc finger fusion proteins that bind to any other contiguous nucleotide sequences of at least 3 nucleotides, or a variant of such a fusion protein.

It is alleged that while the claims are enabled for polydactyl zinc finger fusion proteins that bind the *erbB-2* promoter, they are not enabled for polydactyl zinc finger fusion proteins that bind to any other contiguous nucleotide sequences of at least 3 nucleotides, or a variant of such a fusion protein.

The arguments above, demonstrate that this is patently incorrect. First, it is respectfully submitted that, contrary to the Examiner's assertion, the specification also provides descriptions and working examples of the claimed zinc finger fusion proteins that bind to and/or regulate sequences **other than** the *erbB-2* promoter. As discussed above, Examples 3-6 beginning at page 87 of the specification and Examples 9-11 beginning at page 105 of the specification provide the construction of fusion proteins containing progesterone or estrogen receptor ligand binding domains either fused to zinc finger DNA binding domains alone or to zinc finger domains and transcriptional activators or repressors. The fusion proteins provided in the aforementioned examples were tested for their ability to regulate ***erbB-2*** (Examples 9 and 11), ***integrin β3*** (Examples 9 and 11), **SV40** (Examples 3, 5 and 10) and **CMV** (Example 6) promoters and **various palindromic oligonucleotide target sequences** (Example 4). Further, contrary to the Examiner's assertion, the specification provides not only fusion proteins

containing the amino acid sequence set forth in SEQ ID NO. 1, and fusion proteins containing a zinc finger protein E2C(Sp1), B3B(Sp1) or B3C2(Sp1) and ERD, KRAB, SID **transcription repression domains** that bind the *erbB-2* promoter as acknowledged by the Examiner, but also zinc finger fusion constructs containing **ligand binding domains** such as progesterone or estrogen receptor (*see, e.g.*, Examples 3, 4, 8, 9, 10 and elsewhere throughout the specification) and/or other transcription repression domains and **transcription activation domains** such as VP16 and VP64 (*see, e.g.*, Examples 3, 4, 9, Table 10 beginning at page 106 and elsewhere throughout the specification).

Second, as discussed above, the specification also provides in exquisite detail by its teachings and working examples the types of sequence variants **in general** that retain the properties of a zinc finger binding domain or a transcription regulation domain, and the selection of these variants. Page 33, line 12 to page 50, line 2 of the specification provides in great detail and incorporates by reference what was known to those of skill in the art at the time of filing of the application concerning zinc finger proteins, the modular nature of zinc finger proteins wherein each zinc finger specifically recognizes a 3 nucleotide sequence, the types of zinc finger proteins, specific changes that provide variant zinc finger peptides that retain the characteristics of recognizing zinc finger DNA binding motifs, how to construct, isolate or synthesize such variants, and how to screen for such variants.

The specification further provides numerous working examples and descriptions of the construction and expression of the claimed fusion proteins. Example 1 beginning at page 67 of the specification describes in great detail the construction and testing of exemplary zinc finger domains. Example 1 shows how variants that specifically recognize the "GNN" target triplet nucleotide sequence of a zinc finger modular unit can be selected. Example 1 also provides numerous exemplary target sequences from various regions of the *erbB-2* and *integrin β3* promoters that provide zinc finger recognition motifs. These target sequences (Table 2 of Example 1) along with the exemplary zinc finger modular

units and their recognition sequences provided in Example 1 (Table 1 of Example 1) can be extrapolated to similarly tailor zinc finger binding domains to interact with selected sequences of targets other than the *erbB-2* , *integrin β3* , SV40 and CMV promoter regions and various palindromic oligonucleotide target sequences that are exemplified herein.

Thus, the specification and the working examples enable the construction and tailoring of polydactyl zinc finger fusion proteins to bind to desired recognition motifs of at least three contiguous nucleotides.

Policy Reasoning

Lastly, as the Examiner has acknowledged, the specification provides examples of fusion proteins containing an amino acid sequence as set forth in SEQ ID NO. 1, and fusion proteins containing a zinc finger protein E2C(Sp1), B3B(Sp1) or B3C2(Sp1) and ERD, KRAB, SID domains that bind the *erbB-2* promoter. The specification describes a variety and large number of DBDs, including a numerous zinc finger proteins.

In addition, as described above, the specification provides exemplary constructs of fusion proteins as instantly claimed that modulate *integrin β3* , SV40 and CMV promoter activity. As further discussed above, the specification provides and incorporates by reference descriptions of and rules for the selection of suitable zinc finger fusion protein variants as claimed herein and suitable target sequences in great detail.

It is respectfully submitted that since **(a)** the structural and functional characteristics of the various domain components of the instantly claimed fusion proteins are described in the specification in great detail; **(b)** contrary to the Examiner's assertion and as discussed above, the specification provides working examples and teachings of fusion constructs as instantly claimed that bind to and/or regulate not only the *erbB-2* promoter but also *integrin β3* , SV40 and CMV promoters and various palindromic oligonucleotide sequences; and **(c)** as described above and as acknowledged by the Examiner, several exemplary fusion constructs have been provided, it would be unfair and unduly limiting to

require Applicant to limit these claims to a few exemplary sequences. To do so is contrary to the public policy upon which the U.S. patent laws are based. If Applicant is required to limit the claims to only the exemplified fusion proteins, then those of skill in the art could by virtue of the teachings of this application readily practice what is claimed by substituting other ligand binding domains, zinc finger peptides or modular units thereof, and transcription regulation domains, but avoid infringing such limited claims. To permit that is simply not fair. The instant application exemplifies the means for isolation, modification and screening of domains having the structural and functional limitations as generically claimed, as well as the means for construction and expression of the fusion proteins, and *in vitro* and *in vivo* assays for their sequence specificity and regulatory activity. Having done so, it is now routine to for others to insert other ligand binding, zinc finger, and transcription regulatory domains into the exemplified fusion proteins. Those of skill in the art should not be permitted to make such minor modifications by substitution of a different host and avoid infringing such claims.

Written Description Rejection

Claims 1-3, 5-35, 37-46 and 69-73 are rejected under 35 U.S.C. §112, first paragraph, as allegedly containing subject matter that was not described in the specification in such a way as to reasonably convey to one skilled in the relevant art that the inventor(s) had possession of the claimed subject matter at the time the application was filed. The Examiner alleges that the claims are genus claims drawn to a protein variant of SEQ ID NO. 1, or a fusion protein containing a ligand binding domain derived from an intracellular receptor where the nucleotide binding domain is derived from a zinc finger peptide that binds a sequence of at least 3 nucleotides, or a variant of such a fusion protein. The Examiner further alleges that the specification is not sufficiently descriptive of a "representative number of species" that are representative of the entire genus. It is alleged that while the specification describes fusion proteins containing the amino acid sequence set forth in SEQ ID NO. 1, and fusion proteins containing a

zinc finger protein E2C(Sp1), B3B(Sp1) or B3C2(Sp1) and ERD, KRAB, SID domains that bind the *erbB-2* promoter, there is no description of polydactyl zinc finger fusion proteins that bind to any other contiguous nucleotide sequences of at least 3 nucleotides, or a variant of such a fusion protein.

This rejection is respectfully traversed.

Relevant Law

The purpose behind the written description requirement is to ensure that the patent applicant had possession of the claimed subject mater at the time of filing of the application In re Wertheim, 541 F.2d 257, 262, 191 USPQ 90, 96 (CCPA 1976). The manner in which the specification meets the requirement is not material; it may be met by either an express or an implicit disclosure.

35 U.S.C. §112 requires a written description of the invention. This requirement is distinct from and not coterminous with the enablement requirement:

> The purpose of the 'written description' requirement is broader than to merely explain how to 'make and use'; the applicant must also convey with reasonable clarity to those skilled in the art that, as of the filing date sought, he or she was in possession of the invention. The invention is, for purposes of the 'written description' inquiry, whatever is now claimed." Vas-Cath, Inc. v. Mahurkar, 935 F.2d at 1563-64, 19 USPQ2d at 1117 (emphasis in original).

The issue with respect to 35 U.S.C. §112, first paragraph, adequate written description has been stated as:

> [d]oes the specification convey clearly to those skilled in the art, to whom it is addressed, in any way, the information that appellants invented that specific compound [claimed embodiment] Vas-Cath, Inc. v. Mahurkar, at 1115, quoting In re Ruschig, 390 F.2d 1990, at 995-996, 154 USPQ 118 at 123 (CCPA 1967).

A specification must convey with reasonable clarity to those skilled in the art that, as of the filing date sought, he or she was in possession of the invention, *i.e.*, whatever is now claimed. Vas-Cath, Inc. v. Mahurkar, 935 F.2d 1555, 1563-64, 19 USPQ.2d 1111, 1117 (Fed. Cir. 1991). A written description requirement issue generally involves the question of whether the

subject matter of a claim is supported by or conforms to the disclosure of an application as filed. The test for sufficiency of support in a patent application is whether the disclosure of the application relied upon "reasonably conveys to the artisan that the inventor had possession at that time of the later claimed subject matter." Ralston Purina Co. v. Far-Mar-Co., Inc., 772 F.2d 1570, 1575, 227 USPQ 177, 179 (Fed. Cir. 1985) (quoting In re Kaslow, 707 F.2d 1366, 1375, 217 USPQ 1089, 1096 (Fed. Cir. 1983)) (see also, MPEP 2163.02).

An objective standard for determining compliance with the written description requirement is "does the description clearly allow persons of skill in the art to recognize that he or she invented what is claimed." In re Gosteli, 872 F.2d 1008, 1012, 10 USPQ.2d 1614, 1618 (Fed. Cir.1989).

The Examiner has the initial burden of presenting evidence or reasons why persons skilled in the art would not recognize in an applicant's disclosure a description of the invention defined by the claims. In re Wertheim, 541 F.2d 257, 265, 191 USPQ 90, 98 (CCPA 1976); *See also* Ex parte Sorenson, 3 USPQ.2d 1462, 1463 (Bd. Pat.App. & Inter. 1987). By disclosing in a patent application a device that inherently performs a function or has a property, operates according to a theory or has an advantage, a patent application necessarily discloses that function, theory or advantage, even though it says nothing explicit concerning it. The application may later be amended to recite the function, theory or advantage without introducing prohibited new matter. In re Reynolds, 443 F.2d 384, 170 USPQ 94 (CCPA 1971); and In re Smythe, 480 F. 2d 1376, 178 USPQ 279 (CCPA 1973).

Furthermore, the subject matter of the claims need not be described literally (*i.e.*, using the same terms or *inhaec verba*) in order for the disclosure to satisfy the description requirement. If a claim is amended to include subject matter, limitations, or terminology not present in the application as filed, involving a departure from, addition to, or deletion from the disclosure of the application as filed, the examiner should conclude that the claimed subject matter is not described in that application. This

conclusion will result in the rejection of the claims affected under 35 U.S.C.112, first paragraph - description requirement, or denial of the benefit of the filing date of a previously filed application, as appropriate.

Analysis

First, as discussed above with respect to the enablement rejection and contrary to the Examiner's assertion, the specification also describes zinc finger fusion proteins as claimed that bind to and/or regulate sequences **other than** the *erbB-2* promoter. As discussed above, Examples 3-6 beginning at page 87 of the specification and Examples 9-11 beginning at page 105 of the specification provide the construction of fusion proteins containing progesterone or estrogen receptor ligand binding domains either fused to zinc finger DNA binding domains alone or to zinc finger domains and transcriptional activators or repressors. The fusion proteins provided in the aforementioned examples were tested for their ability to regulate ***erbB-2*** (Examples 9 and 11), ***integrin β3*** (Examples 9 and 11), **SV40** (Examples 3, 5 and 10) and **CMV** (Example 6) promoters and **various palindromic oligonucleotide target sequences** (Example 4).

Further, the specification describes not only fusion proteins containing the amino acid sequence set forth in SEQ ID NO. 1, and fusion proteins containing a zinc finger protein E2C(Sp1), B3B(Sp1) or B3C2(Sp1) and ERD, KRAB, SID **transcription repression domains** that bind the *erbB-2* promoter as acknowledged by the Examiner, but also zinc finger fusion constructs containing **ligand binding domains** such as progesterone or estrogen receptor (*see, e.g.*, Examples 3, 4, 8, 9, 10 and elsewhere throughout the specification) and/or other transcription repression domains and **transcription activation domains** such as VP16 and VP64 (*see, e.g.*, Examples 3, 4, 9, Table 10 beginning at page 106 and elsewhere throughout the specification).

The passages reproduced above, demonstrate clear posssession and appreciation of a large number of DBDs. As noted above, the specification

provides similarly detailed guidance for selection/synthesis and/or identification of the other domains in fusion proteins as claimed. The instant specification provides fusion proteins containing zinc finger peptide nucleotide binding domains operatively linked to ligand binding domains from an intracellular receptor that can further comprise an operatively linked transcription regulating domain. The specification provides extensive lists of examples of each domain and exemplifies preparation of various fusion proteins having the structural and functional limitations as claimed. There is nothing of record that suggests Applicant did not contemplate or appreciate a fusion protein containing a ligand binding domain from an intracellular receptor, a DNA binding domain that is a zinc finger peptide assembled from modular units that specifically recognize 3 nucleotide sequences, and a transcription regulation domain. On the contrary, the specification clearly evidences that applicant contemplated generic subject matter. Moreover, the specification clearly sets forth the types of variants in each domain that would satisfy the structural and functional limitations of the claims.

The specification exemplifies in great detail the construction and expression of the claimed fusion genes, assays that screen for variants that fall within the scope of the claims by measuring specific binding of each of the domains, and assays to measure their potential as regulators of gene expression. As discussed above, the classes of molecules belonging to each of the domains of the instantly claimed fusion proteins had been characterized in exquisite detail in the art as of the effective filing date of the instant application and have also been extensively described in the specification so as to be adequately descriptive of "variants" that fall within the scope of the claims. Furthermore, as is well known to those of skill in the art (*see* above), and as described extensively in the specification, recombinant technology and binding assays may be uniformly applied to any or all of the molecules described in the specification due to their common structural and functional characteristics such as their target recognition sites, the structural motifs that create the specificity

of recognition, and methods by which their ligand binding characteristics may be altered.

The ligand binding domain and variants thereof that fall within the scope of the claims are described extensively at, for example, page 32, line 7 to page 33, line 11 of the specification, which provides methods, known to those of skill in the art, to prepare and characterize variants of the ligand binding domain, including specific changes that will provide altered endogenous or exogenous ligand specificity as desired. Page 33, line 12 to page 50, line 2 of the specification provides in exquisite detail and incorporates by reference what was known to those of skill in the art at the time of filing of the application concerning zinc finger proteins, the modular nature of zinc finger proteins wherein each zinc finger specifically recognizes a 3 nucleotide sequence, the types of zinc finger proteins, specific changes that provide variant zinc finger peptides that retain the characteristics of recognizing zinc finger DNA binding motifs, the rules for constructing, isolating or synthesizing such variants, and how to screen for such variants. Page 50, line 3 to page 52, line 5 of the specification describes known transcriptional regulatory domains and selection and modifications thereof. At, for example, page 31, line 6, to page 32, line 6, the specification describes how to construct the claimed fusion proteins from the various domains and their variants. In addition, numerous working examples, discussed above, are provided throughout the specification, as are exemplary fusion proteins, encoded by SEQ ID NOS. 1-18. As described above, the working examples also set forth in great detail the construction and screening of variants that fall within the scope of the instant claims

For example, Page 32, line 7 to page 33, line 11 of the specification states:

> **1. Ligand Binding Domain (LBD)**
>
> The ligand binding domain is derived from an intracellular receptor, and is preferably derived from a nuclear hormone receptor. The LBD of

an intracellular receptor includes the approximately 300 amino acids from the carboxy terminal, which can be used with or without modification.

By mutation of a small number of residues ligand specificity can be altered. The ligand binding domain can be modified, such as by truncation or point mutation to alter its ligand specificity permitting gene regulation by non-natural or non-native ligands.

Exemplary hormone receptors are steroid receptors, which are well known in the art. Exemplary and preferred steroid receptors include estrogen and progesterone receptors and variants thereof. Of particular interest are ligand binding domains that exhibit altered ligand specificity so that the LBD does not respond to the natural hormone, but rather to a drug, such as RU486, or other inducer. **Means to modify and test the specificity of ligand binding domains and to identify ligands therefor are known (see, U.S. Patent No. 5,874,534; U.S. Patent No. 5,935,934; and International PCT application No. 98/18925, which is based on U.S. provisional application Serial No. 60/029,964; International PCT application No. 96/40911, which is based on U.S. application Serial No. 08/479,913).**

The LBD can be modified by deletion of from about 1 up to about 150, typically 120, amino acids on the carboxyl terminal end of the receptor from which the LBD derives. Systematic deletion of amino acids and subsequent testing of the ligand specificity and of the resulting LBD can be used to empirically identify mutations that lead to modified LBDs that have desired properties, such as preferential interaction with non-natural ligands. Exemplary mutations are described in the Examples herein, and also are known to those of skill in the art (see, *e.g.*, U.S. Patent No. 5,874,534; U.S. Patent No. 5,935,934; U.S. Patent No. 5,364,791; and International PCT application No. 98/18925, which is based on U.S. provisional application Serial No. 60/029,964; International PCT application No. 96/40911, which is based on U.S. application Serial No. 08/479,913) and references cited therein. **Hence a LBD or modified form thereof prepared by known methods is obtained and operably linked to a DBD; a TRD is also linked as needed.** (emphasis added).

For example, at page 33, lines 15-29 of the specification:

Zinc fingers are ubiquitous proteins, and many are well-characterized. **For example, methods and rules for preparation and selection of zinc fingers based upon the C2H2 class of zinc fingers with unique specificity are known** (see, *e.g.*, International PCT application No. WO 98/54311 and International PCT application No. 95/19431; see, also U.S. Patent No. 5,789,538; Beerli *et al.* (1999) *Proc. Natl. Acad. Sci. U.S.A. 96*:2758-

2763; Beerli *et al.* (1995) *Proc. Natl. Acad. Sci. U.S.A. 95*:14628-14633; see, also U.S. application Serial No. 09/173,941, filed 16 October, 1998, published as International PCT application No. WO 00/23464). Exemplary targeting sequences are provided herein.

Furthermore, other zinc fingers can be similarly identified and the rules known for the C2H2 can be applied to modification of the specificity of such zinc fingers or alternative rules unique to each class can be deduced in a similar manner. (emphasis added).

Further, *e.g.*, at page 34, lines 16-30:

For example, zinc finger variants can be prepared by identifying a zinc finger or modular unit thereof, creating an expression library, such as a phage display library (see, *e.g.*, International PCT application No. WO 98/54311, Barbas *et al.* (1991) *Methods 2*:119; Barbas *et al.* (1992) *Proc. Natl. Acad. Sci. U.S.A. 89*:4457), encoding polypeptide variants of the zinc finger or modular unit thereof, expressing the library in a host and screening for variant peptides having a desired specificity. Zinc fingers may also be constructed by combining amino acids (or encoding nucleic acids) according to the known rules of binding specificity and, if necessary, testing or screening the resulting peptides to ensure the peptide has a desired specificity. **Because of the modular nature of zinc fingers, where each module can be prepared to bind to three nucleotide sequence, peptides of any specificity can be prepared from the modules.** The number of modules used depends upon the specificity of gene targeting desired. (emphasis added).

Example 1 beginning at page 67 of the specification provides a number of zinc finger variants produced in the manner described throughout the specification.

Further, for example, at page 50, lines 13-18 of the specification:

Selection of the TRD

Transcription regulating domains are well known in the art. Exemplary and preferred transcription repressor domains are ERD, KRAB, SID, Deacetylase, and derivatives, multimers and combinations thereof such as KRAB-ERD, SID-ERD, $(KRAB)_2$, $(KRAB)_3$, KRAB-A, $(KRAB\text{-}A)_2$, $(SID)_2$ (KRAB-A)-SID and SID-(KRAB-A).

Transcription activator domains are also described in the specification, for example, at page 7, lines 10-13 and at page 52, lines 1-5, which provide

examples of transcription activation domains that include VP16, VP64, TA2, STAT-6, p65, relA and derivatives, multimers and combinations thereof.

Based on the knowledge of those of skill in the art at the time that the application was filed (*see* descriptions above), the specification fully describes the subject matter as claimed at the time that the application was filed. Furthermore, it must be noted that the Examples set forth standard technologies and specific structural features for the construction and assay of the fusion protein variants that fall within the scope of the claims.

Therefore, because there is extensive written description as to the identity, structural features, sequence variants, and screens to identify the sequence variants of the claimed fusion proteins, Applicant had possession of the claimed subject matter at the time of filing of the application.

THE REJECTION OF CLAIMS 1-35, 37-46 and 69-73 UNDER 35 U.S.C. §112, SECOND PARAGRAPH

Claims 1-35, 37-46 and 69-73 are rejected under 35 U.S.C. 112, second paragraph, as allegedly being indefinite for failing to particularly point out and distinctly claim the subject matter that applicants regard as the invention. Various bases for this rejection are set forth and each is discussed in turn. Reconsideration of the grounds for rejection is respectfully requested in view of the amendments of the claims and the following remarks.

Relevant Law

Definiteness of claim language must be analyzed, not in a vacuum, but in light of (1) the content of the particular application disclosure, (2) the teachings of prior art, and (3) the interpretation claims would be given by one possessing the ordinary level of skill in the pertinent art at the time the invention was made. Claims need only "reasonably apprise those skilled in the art" of their scope and be "as precise as the subject permits." Hybritech Inc. v. Monoclonal Antibodies, Inc., 231 USPQ 81, 94 (Fed. Cir. 1986), cert. den., 480 U.S. 947 (1987). The Court in Orthokinetics, Inc v. Safety Travel Chairs, Inc., 1 USPQ2d 1081 (Fed. Cir. 1986) held that a claim limitation requiring that a pediatric

wheelchair part be "so dimensioned as to be insertable through the space between the doorframe of an automobile and one of the seats" is definite. The Court stated:

> The phrase 'so dimensioned' is as accurate as the subject matter permits, automobiles being of various sizes. As long as those of ordinary skill in the art realized that the dimensions could be easily obtained, § 112, 2d ¶ requires nothing more. The patent law does not require that all possible lengths corresponding to the spaces in hundreds of different automobiles be listed in the patent, let alone that they be listed in the claims.

1 USPQ2d at 1088.

When one skilled in the art would understand all of the language in the claims when read in light of the specification, a claim is not indefinite.

Applicant is unaware of any requirement that terms be defined in the claims when one of skill in the art can readily determine the meaning of the term based on the description and definitions provided in the specification. In this respect, applicant is entitled to be its own lexicographer [see, e.g., MPEP 2111.01 "Applicant may be his or her own lexicographer as long as the meaning assigned to the term is not repugnant to the term's well known usage and utilize terms within the claims that are clear from a reading of the specification]. In re Hill, 73 USPQ 482 (CCPA 1947)". When applicant has provided definitions in the specification, the claims are interpreted in light of such definition.

35 U.S.C. §112, second paragraph requires only reasonable precision in delineating the bounds of the claimed invention. The claim language is satisfactory if it reasonably apprises those of skill in the art of the bounds of the claimed invention and is as precise as the subject matter permits. Shatterproof Glass Corp.v. Libby-Owens Ford Col, 758 F.2d 613, 624, 225 USPQ 634, 641 (Fed. Cir), cert dismissed, 106 S. Ct. 340 (1985).

The amount of detail required to be included in the claims depends on the particular invention and the prior art and is not to be viewed in the abstract, but in conjunction with whether the specification is in compliance with the first

paragraph of 35 U.S.C. §112. If the claims, read in light of the specification, reasonably apprise those skilled in the art of the utilization and scope of the invention, and if the language is as precise as the subject matter permits, the courts can demand no more:

> [i]t is not necessary that a claim recite each and every element needed for the practical utilization of the claimed subject matter (Bendix Corp. v United States, 600 F.2d 1364, 1369, 220 Ct. Cl. 507,514, 204 USPQ 617, 621 (1979); See, also, Carl Zeiss Stiftung v. Renishaw plc, 20 USPQ2d 1094, 1101).

Analysis

1) The Examiner alleges that claims 1, 8, 21, 22, 25, 32, 39, 43 and 45 are indefinite in their recitation of the term "derived from." It is alleged that it is unclear whether the term imposes a required limitation on the claim such that it only encompasses, for example, polynucleotides encoding the receptor, or only sequences produced by digestion with restriction enzymes of DNA isolated from human tissue that contains polynucleotides encoding the receptor, if the claim encompasses all polynucleotide sequences encoding the receptor. This rejection is respectfully traversed.

It is clear from the disclosure in the specification and the knowledge of one of skill in the art that "derived from" indicates the source from which the substance is obtained. References to sources from which components are "derived" abound throughout the instant specification and in the art. "Derived from" is a term of art that is commonly used in instances such as the construction of multi-component biological products including fusion proteins.

With respect to the disclosure in the specification, for example, page 15, lines 6-11 of the specification recites that the ligand binding domain is "derived from" the 300 amino acid carboxyl terminal half of intracellular receptors, and is the portion of the receptor protein with which a ligand interacts. Thus, the source of the ligand binding domain is the intracellular receptor. As the specification, for example, at page 3, lines 7-13 and the claims (*e.g.*, Claim 1) further provide, this ligand binding domain, *i.e.*, the portion with which the

ligand interacts, may be modified to obtain altered ligand specificity. The ligand binding domains themselves are described and characterized in great detail (*see, e.g.*, Page 32, line 7 to page 33, line 11 of the specification) so that the metes and bounds of what constitutes this domain that is "derived from" an intracellular receptor is clear. The metes and bounds of the phrase "derived from" is further adequately established because the meaning of the phrase in the context of obtaining a component of a fusion protein as claimed herein is known in light of the disclosure in the specification and the knowledge of one of skill in the art.

The phrase "derived from" is used as a common term of art to denote sources of domains of fusion proteins. For example, a search of the USPTO database in just the year preceding the earliest priority date of the instant application for patents containing a combination of the terms "receptor" and "derived from" in the claims yielded several hundred patents claiming fusion proteins and indicating where the components (domains) of the fusion proteins are "derived from." Some of these patents are set forth below:

1) U.S. Patent No. 5,843,446, claims compositions containing a peptide with an invasin domain that is "derived from" variant strains of *Yersinia spp.*;

2) U.S. Patent No. 5,798,229, claims a bispecific antibody fragment containing an antibody determinant that is specific for an epitope on a tumor cell where the determinant is "derived from" various monoclonal antibodies;

3) U.S. Patent No. 5,783,420, claims a method of regulating gene expression by introducing an expression vector encoding antibody derivatives that bind to *cis*-regulatory DNA sequences. The C-terminal region of the antibody derivatives is claimed as containing a DNA-binding domain "derived from" a protein capable of specifically binding to a *cis*-regulatory DNA element.

4) U.S. Patent No. 5,783,405, claims a method for ascertaining modulation of intracellular signal transduction by providing a receptor for activated protein kinase C (PKC) or a peptide "derived from" PKC that binds specifically to catalytically active PKC via a noncatalytic site of PKC;

5) U.S. Patent No. 5,843,728, claims a membrane-bound chimeric receptor protein that has an intracellular portion that is "derived from" a T cell receptor, a B cell receptor or an Fc receptor;

6) U.S. Patent No. 5,858,657, claims a method of producing specific binding pair members containing first and second polypeptide chains where either or both of the chains are "derived from" various sets of immunoglobulin genes;

7) U.S. Patent No. 5,856,456, claims a DNA molecule encoding a fusion polypeptide containing first and second polypeptides that are "derived from" members of the immunoglobulin superfamily;

8) U.S. Patent No. 5,159,194 claims a method for the mass spectrometric "analysis" of a sample in which the ionization steps are defined;

9) U.S. Patent No. 5,912,170, claims an isolated cytotoxic T cell expressing at least two membrane-bound chimeric receptors where one of the receptors contains an intracellular portion that is "derived from" CD28;

10) U.S. Patent No. 5,939,531, claims a fusion protein containing a single-chain recombinant antibody where the heavy chain and light chain variable domains are "derived from" various specified mouse monoclonal antibodies;

11) U.S. Patent No. 5,955,300, claims a method for producing a soluble polypeptide fraction by inserting a DNA molecule containing a fusion of fragments of cDNA coding for the polypeptide regions corresponding to LAG-3 or the cDNA coding for the portion of the immunoglobulin that is "derived from" LAG-3 into a suitable host expression system;

12) U.S. Patent No. 5,958,706, claims a protein-bound magnetic particle produced in the cell of a magnetic bacterium that contains a membrane bound portion that is "derived from" the mps protein or the mag A protein;

and many others.

Thus, claims 1, 8, 21, 22, 25, 32, 39, 43 and 45 and its dependent claims are not indefinite because in light of the disclosure in the specification and what is known to one of skill in the art, "derived from" is accepted as a

term of art indicating the source from which a substance/component is obtained. In an effort to advance this case to issue, "derived from" is amended to recite "from." There, however, is not intent to change the intened meaning of the language of the claim as discussed above.

2) The Examiner alleges that Claim 13 is unclear in its recitation of the term "selectivity".

As discussed previously, there is basis in the specification for the term selectivity as meaning an altered binding specificity relative to the native ligand binding domain, and methods of assaying for this selective preference is incorporated by reference in the specification at page 5, lines 5-20 (see, *e.g.*, U.S. Patent No. 5,874,534 and Wang *et al.* (1994) *Proc. Natl. Acad. Sci. U.S.A. 91*:8180-8184) and in Example 14 at page 113 of the specification. Nonetheless, this rejection is obviated by amending Claim 13 to recite that the receptor variant has altered ligand specificity for endogenous and exogenous ligands relative to the native receptor. Basis for this amendment may be found in the specification, *e.g.*, at page 5, lines 10-20; at page 15, lines 3-16; at page 31, lines 22-26; and at page 32, line 7 to page 33, line 11.

3) The Examiner alleges that claim 5 is indefinite in its recitation of the term "substantially". Responsive to Applicant's amendment of Claim 5 to recite "relative to exogenous or non-natural ligands", it is alleged that this phrase does not clarify the metes and bound of the claim because no specific parameters considered to define "substantially" have been added. This rejection is respectfully traversed.

The amount of detail required to be included in the claims depends on the particular subject matter and the prior art and is not to be viewed in the abstract, but in conjunction with whether the specification is in compliance with the first paragraph of 35 U.S.C. §112. If the claims, read in light of the specification, reasonably apprise those skilled in the art of the utilization and scope of the invention, and if the language is as precise as the subject matter permits, the courts can demand no more:

> [i]t is not necessary that a claim recite each and every element needed for the practical utilization of the claimed subject matter (Bendix Corp. v United States, 600 F.2d 1364, 1369, 220 Ct. Cl. 507,514, 204 USPQ 617, 621 (1979); See, also, Carl Zeiss Stiftung v. Renishaw plc, 20 USPQ2d 1094, 1101).

Recently, it has been ruled that terms such as "about" and "substantially" are descriptive and can be used in place of a strict numerical boundary for a claim parameter when the nature of the subject matter permits such a description.

> Expressions such as 'substantially' are used in patent documents when warranted by the nature of the invention, in order to accommodate the minor variations that may be appropriate to secure the invention (Verve, LLC v. Crane Cams, Inc., 311 F.3d 1116 (2002)).

The nature of the subject matter of Claim 5 is such that precise parameters in support of "substantially" are not required. Claim 5 provides relative activation of the modified ligand binding domain for endogenous versus exogenous or non-natural ligands such that activation by exogenous or non-natural ligands is greater than activation by endogenous ligands. It is enough that there is a preference for exogenous or non-natural ligands as demonstrated by the methods of measurement that are provided in the specification (*see, e.g.*, Example 14 beginning at page 113) to lend clarity to the metes and bound of the claimed subject matter. Moreover, none of the references cited in connection with examination of the instant claims provides fusion proteins containing nucleotide binding domains operatively linked to ligand binding domains where the ligand binding domain is modified such that it shows a preference for non-natural or exogenous ligands rather than endogenous ligands. In the context of the instantly claimed subject matter it is, therefore, sufficient to recite that the modified ligand binding domain is not substantially activated by endogenous ligands relative to exogenous or non-natural ligands to determine the metes and bounds of the claim.

4) The Examiner alleges that claims 1, 8, 20, 22, 25, 32, 39, 43, 45, 69 and 73 recite the term "specifically", which is indefinite. Claims 2, 3, 5-7, 9-19, 21, 23, 24, 26-31, 33-38, 40-42, 44, 46 and 70-72 are also rejected insofar as they depend from the aforementioned claims. It is alleged that it is not clear to what degree the nucleotide binding domain must interact with the target nucleotides in order to meet the limitation of "specifically."

Sequence-specificity of DNA or RNA-binding proteins for their cognate nucleic acid sequences is a term of art that is known to refer to sequence selectivity that is achieved by the ability of specific nucleic acid sequences to adopt a structure required for binding to a particular protein at lower free energy cost relative to other sequences. For example, Steitz, "*Structural Studies of Protein-Nucleic Acid Interactions*" (Cambridge University Press, 1990), provides an overview of studies on sequence-specific binding by proteins that provide factors influencing a preference of nucleic acid binding proteins for one sequence over any others (*see, e.g.*, Steitz at pp. 4-10, attached hereto). It is understood by those of skill in the art that sequence-specificity indicates the recognition of certain sequences over others by nucleic acid binding proteins, without the provision of specific parameters that measure the extent of such interaction (*see, e.g,* Steitz). The rejected claims clearly recite the sequences with which the claimed DNA-binding fusion proteins interact, and such sequence identification is sufficient to indicate the sequence specificity of the protein. In addition, such claims recite that the fusion protein is a "is a ligand activated transcriptional regulator," thereby further describing the nature of the interaction.

Further, the specification at *e.g.*, page 33, line 12 to page 34, line 15 provides basis, supported by publications incorporated by reference, for determining the specificity of binding of modular units of a zinc finger peptide that specifically interact with 3 nucleotide contiguous sequences. Also, *e.g.*, page 39, lines 17-27 of the specification, and Example 1 at page 67 of the

specification provide in great detail how specific binding of zinc finger peptides to their recognition motifs is determined.

Notwithstanding the above, the claims have been amended to remove the word "specifically." The metes and bounds of the claims as amended are clear because zinc finger peptides are known to have sequence specificity for contiguous sequences of at least about 3 nucleotides (*see* specification, *e.g.*, at page 5, line 21 to page 6, line 2; page 38, lines 16-24). Further, the claims inherently provide the standard for specificity as described in the specification (*see* specification, *e.g.*, at page 10, lines 26-29) by reciting that the zinc finger polypeptide or modular portion thereof that is a component of the fusion protein interacts with a contiguous sequence of at least about 3 nucleotides so that the fusion protein is a ligand activated transcriptional regulator.

5) The Examiner alleges that Claim 21 is vague and indefinite in its recitation of the terms "KRAB-ERD", "SID-ERD", "$(KRAB)_2$", *etc.* This rejection is respectfully traversed.

The specification at *e.g.*, page 14, line 26 to page 15, line 1 and page 51, lines 3-30 provides and incorporates by reference the published sequences and descriptions of the proteins represented as ERF repressor domain (ERD) (Sgouras *et al.* (1995) *EMBO J. 14*:4781-4793), defined by amino acids 473 to 530 of the *ets2* repressor factor (ERF) and sequence set forth in Figure 1 at page 4783); Krüppel-associated box (KRAB) domain (Margolin *et al.* (1994) *Proc. Natl. Acad. Sci. USA 91*:4509-4513; sequence set forth in Figure 1 at page 4510); mSIN3 interaction domain (SID) (Ayer *et al.* (1996) *Mol. Cell. Biol. 16*:5772-5781; sequence set forth in Figure 1 at page 5774); Deacetylase, and derivatives, multimers and combinations thereof (*see also*, *e.g.*, page 7, lines 24-29 of the specification). Therefore, it is respectfully submitted that Claim 21 is not indefinite.

THE REJECTION OF CLAIMS 1-3, 5-35, 37, 38, 43-46 and 69-73 UNDER 35 U.S.C. § 102

Claims 1-3, 5-35, 37, 38, 43-46 and 69-73 are rejected under 35 U.S.C. §102(a) as being anticipated by Beerli *et al.* (*Proc. Natl. Acad. Sci. U.S.A.95*:14628-14633 (1998)). It is alleged that Beerli *et al.* discloses using a family of zinc finger domains that recognize sequences of the 5'-GNN-3'motif to construct polydactyl zinc finger - transcription factor fusion proteins that specifically recognize 9 or 18 bp sequences. It is further alleged that the transcription activator or repressor domains of the fusion proteins disclosed in Beerli *et al.* include VP16, VP64, KRAB, ERD and SID. Furthermore, Beerli *et al.* is alleged to disclose the achievement of gene-specific activation or repression using the zinc finger - transcriptional factor fusion proteins, as well as binding to a single site in native *erbB-2* promoter. The Examiner concludes that Beerli *et al.* anticipates the claims.

This rejection is respectfully traversed.

Relevant law

Anticipation requires the disclosure in a single prior art reference of each element of the claim under consideration. In re Spada, 15 USPQ2d 1655 (Fed. Cir, 1990), In re Bond, 15 USPQ 1566 (Fed. Cir. 1990), Soundscriber Corp. v. U.S.. 360 F.2d 954, 148 USPQ 298, 301, adopted 149 USPQ 640 (Ct. Cl.) 1966. See, also, Richardson v. Suzuki Motor Co., 868 F.2d 1226, 1236, 9 USPQ2d 1913,1920 (Fed. Cir.), cert. denied, 110 S.Ct. 154 (1989). "[A]ll limitations in the claims must be found in the reference, since the claims measure the invention". In re Lang, 644 F.2d 856, 862, 209 USPQ 288, 293 (CCPA 1981). Moreover it is incumbent on Examiner to identify wherein each and every facet of the claimed invention is disclosed in the reference. Lindemann Maschinen-fabrik Gmbh v. American Hoist and Derrick Co., 730 F.2d 1452, 221 USPQ 481 (Fed. Cir. 1984).

Further, the reference must describe the claimed subject matter sufficiently to have placed a person of ordinary skill in the art in possession of

the invention. Prior art does not anticipate a thing or process unless it is enabling; an anticipatory publication must describe the claimed invention with sufficient clarity and specificity so that one skilled in the relevant art could practice the subject matter of the patent without assistance from the patent claimed to have been anticipated Columbia Broadcasting System v. Sylvania Elec. Products, Inc., 415 F.2d 719, 735, 162 USPQ 577 (1st Cir.1968) cert. denied, 396 U.S. 1061, 164 USPQ 321 (1970).

"Before any publication can amount to a statutory bar to the grant of a patent, its disclosure must be such that a skilled artisan could take its teachings in combination with his own knowledge of the particular art and be in possession of the invention." Titanium Metals crp. v Mossinghoff, 603 F.Supp. 87,0, 225 USPQ 673 (1984) quoting In re Application of Le Grice 49 CCPA 1124, 301 F.2d 9333

The claims

Claim 1 is directed to a fusion protein containing a nucleotide binding domain operatively linked to a ligand binding domain from an intracellular receptor. In the elected species the nucleotide binding domain is a polydactyl zinc finger peptide or modular portion thereof that specifically interacts with a contiguous nucleotide sequence of at least about 3 nucleotides; the ligand binding domain is modified to change its ligand specificity compared to the native hormone receptor; and the fusion protein is a ligand activated transcriptional regulator. Claims 2, 3, 5-7 and 9-19, 23-24, 26-31, 33 and 69-73 are all directed to Claim 1 or dependents thereof that specify features of elements of Claim 1 or dependents thereof, or add further elements, such as a transcription regulating domain (Claim 2), or are directed to nucleic acid molecules encoding the fusion proteins of Claims 1 or 2, or to vectors containing the sequence of nucleic acid molecules encoding the fusion proteins of Claims 1 or 2, to cells containing these vectors, or to non-viral delivery systems containing the fusion protein of Claim 1.

Claim 8 is directed to a fusion protein containing a nucleotide binding domain operatively linked to a ligand binding domain from an intracellular receptor, where the nucleotide binding domain is a polydactyl zinc finger peptide or modular portion thereof that specifically interacts with a contiguous nucleotide sequence of at least about 3 nucleotides; the zinc finger peptide contains C2H2 zinc-finger modular units that specifically interacts with a contiguous nucleotide sequence of at least about 3 nucleotides; and the fusion protein is a ligand activated transcriptional regulator.

Claim 20 is directed to a fusion protein containing a nucleotide binding domain operatively linked to a transcription regulating domain and a ligand binding domain from an intracellular receptor where the nucleotide binding domain is a polydactyl zinc-finger or a modular portion thereof that specifically interacts with a contiguous nucleotide sequence of at least about 3 nucleotides; the transcription regulating domain contains a transcription repressor and the fusion protein is a ligand activated transcriptional regulator. Claim 21 specifies the types of transcription repressors in the fusion protein of Claim 20.

Claim 22 is directed to a fusion protein containing a nucleotide binding domain operatively linked to a transcription regulating domain and a ligand binding domain from an intracellular receptor where the nucleotide binding domain is a polydactyl zinc-finger or a modular portion thereof that specifically interacts with a contiguous nucleotide sequence of at least about 3 nucleotides; the fusion protein is a ligand activated transcriptional regulator; and the fusion protein is encoded by the sequence of nucleotides set forth in any of SEQ ID NOS. 1-18.

Claim 25 is directed to a nucleic acid molecule containing a sequence of nucleotides encoding a fusion protein where the fusion protein contains a nucleotide binding domain operatively linked to a transcription regulating domain and a ligand binding domain from an intracellular receptor and the nucleotide binding domain is a polydactyl zinc-finger or a modular portion thereof that specifically interacts with a contiguous nucleotide sequence of at least about 3

nucleotides; the fusion protein is a ligand activated transcriptional regulator; and the fusion protein is encoded by the sequence of nucleotides set forth in SEQ ID NO. 1.

Claim 32 is directed to a viral vector containing a sequence of nucleotides encoding a fusion protein, where the fusion protein contains a nucleotide binding domain operatively linked to a ligand binding domain from an intracellular receptor and the nucleotide binding domain is a polydactyl zinc finger peptide or modular portion thereof that specifically interacts with a contiguous nucleotide sequence of at least about 3 nucleotides; and the fusion protein is a ligand activated transcriptional regulator. Dependent claims 34, 35, 37 and 38 specify particular viruses from which the viral vector of Claim 32 is derived.

Claim 43 is directed to a composition for regulating gene expression that includes an effective amount of a fusion protein contains a nucleotide binding domain operatively linked to a ligand binding domain from an intracellular receptor where the nucleotide binding domain is a polydactyl zinc finger peptide or modular portion thereof that specifically interacts with a contiguous nucleotide sequence of at least about 3 nucleotides and the fusion protein is a ligand activated transcriptional regulator; or a nucleic acid molecule containing a sequence that encodes the fusion protein and a pharmaceutically acceptable excipient. Dependent Claim 44 specifies that the composition of Claim 43 is formulated for a single dosage administration.

Claim 45 is directed to a composition for regulating gene expression that includes an effective amount of a fusion protein contains a nucleotide binding domain operatively linked to a ligand binding domain from an intracellular receptor where the nucleotide binding domain is a polydactyl zinc finger peptide or modular portion thereof that specifically interacts with a contiguous nucleotide sequence of at least about 3 nucleotides and the fusion protein is a ligand activated transcriptional regulator; and a pharmaceutically acceptable excipient.

Claim 46 specifies the number of response elements in the regulatable expression cassette that is a component of the combination of Claim 39, from which it depends. Claim 39 is directed to a combination that includes a fusion protein containing a nucleotide binding domain operatively linked to a ligand binding domain from an intracellular receptor where the nucleotide binding domain is a polydactyl zinc finger peptide or modular portion thereof that specifically interacts with a contiguous nucleotide sequence of at least about 3 nucleotides and the fusion protein is a ligand activated transcriptional regulator; or a nucleic acid molecule containing a sequence of nucleotides that encodes the fusion protein; and a regulatable expression cassette that contains at least one response element recognized by the nucleic acid binding domain of the fusion protein.

Thus, all the claims contain as elements (1) a fusion protein containing a nucleotide binding domain operatively linked to a ligand binding domain; and (2) a fusion protein that is a ligand activated transcriptional regulator.

Differences between the disclosure of Beerli et al. and the claimed subject matter

Beerli *et al.* is directed to the construction of polydactyl zinc finger proteins that specifically bind to 9 or 18 bp target nucleotide sequences in a gene-specific manner. Beerli *et al.* further discloses fusion of the polydactyl zinc finger proteins with KRAB, SID or ERD repressor domains, or VP16 or VP64 activation domains.

Beerli *et al.* does not disclose fusion proteins containing ligand binding domains. The fusion proteins disclosed in Beerli *et al.* contain nucleotide binding zinc finger domains and transcription activation or transcription repression domains, but not ligand binding domains.

As disclosed and claimed in the instant application, the ligand binding domains are the ligand binding domains from intracellular receptors such as, for example, the estrogen or progesterone receptors. Beerli *et al.* also does not disclose ligand activated transcriptional regulation. Therefore, Beerli *et al.* does

not anticipate the claims, all of which include as elements (1) a fusion protein containing a nucleotide binding domain operatively linked to a ligand binding domain; and (2) a fusion protein that is a ligand activated transcriptional regulator. Since anticipation requires that a reference disclose all elements as claimed, Beerli *et al.* does not anticipate any of the claims.

THE REJECTION OF CLAIMS 1-3, 5-35, 37-46 and 69-73 UNDER 35 U.S.C. § 103

Claim 1-3, 5-35, 37-46 and 69-73 are rejected under 35 U.S.C. § 103(a) as allegedly being obvious over the teachings of Beerli *et al.* It is alleged that Beerli *et al.* teaches a "general strategy" for the production of gene switches by using a family of zinc finger domains that recognize sequences of the 5'-GNN-3'motif to construct polydactyl zinc finger - transcription factor fusion proteins that specifically recognize 9 or 18 bp sequences, where the transcription factors include transcriptional activators VP16 and VP64; and transcriptional repressors KRAB, ERD and SID. It is further alleged that Beerli *et al.* teaches that the fusion proteins generated according to the methods provided therein are potent transcription factors that control gene activation and repression. Furthermore, the Examiner alleges that Beerli *et al.* teaches potential utility of the fusion proteins in diagnostic or therapeutic applications. The Examiner concludes that it would have been *prima facie* obvious to one of skill in the art to use the nucleic acids encoding the fusion proteins of Beerli *et al.* in a method of gene therapy, and that particular motivation is allegedly provided at page 14633, which states that "These proteins might also be used in gene therapy applications to inhibit the production of viral gene products or to activate genes involved in fighting disease."

This rejection is respectfully traversed.

Relevant law

In order to set forth a *prima facie* case of obviousness under 35 U.S.C. § 103: (1) there must be some teaching, suggestion or incentive supporting the combination of cited references to produce the claimed invention (*ACS Hospital Systems, Inc. v. Montefiore Hospital*, 732 F.2d 1572, 1577, 221 USPQ 929, 933 (Fed. Cir. 1984)) and (2) the combination of the cited references must actually teach or suggest the claimed subject matter. Further, that which is within the capabilities of one skilled in the art is not synonymous with that which is obvious. *Ex parte Gerlach*, 212 USPQ 471 (Bd. APP. 1980). Obviousness is tested by "what the combined teachings of the references would have suggested to those of ordinary skill in the art" *In re Keller*, 642 F.2d 413, 425, 208 USPQ 871, 881 (CCPA 1981), but it cannot be established by combining the teachings of the prior art to produce the claimed subject matter, absent some teaching or suggestion supporting the combination (*ACS Hosp. Systems, Inc. v Montefiore Hosp.* 732 F.2d 1572, 1577. 221 USPQ 929, 933 (Fed. Cir. 1984)). "To imbue one of ordinary skill in the art with knowledge of the invention in suit, when no prior art reference or references of record convey or suggest that knowledge, is to fall victim to the insidious effect of a hindsight syndrome wherein that which only the inventor taught is used against its teacher" *W.L. Gore & Associates, Inc. v Garlock Inc.*, 721 F.2d 1540, 1553, 220 USPQ 303, 312-13 (Fed. Cir. 1983).

The prior art must provide a motivation whereby one of ordinary skill in the art would have been led to do that which the applicant has done. *Stratoflex Inc. v Aeroquip Corp.*, 713 F.2d 1530, 1535, 218 USPQ 871, 876 (Fed. Cir. 1983). In addition, the mere fact that the prior art may be modified in the manner suggested by the Examiner does not make the modification obvious unless the prior art suggests the desirability of the modification. *In re Fritch*, 23 USPQ 1783 (Fed. Cir. 1992).

Also, it is impermissible to ignore the advantages, properties, utilities and unexpected results that flow from the claimed invention; they are part of the invention as a whole. *In re Sernaker*, 702 F.2d 989, 217 USPQ 1 (Fed. Cir. 1983). Unexpected properties must always be considered when determining obviousness. A compound's structure and properties are inseparable so that unexpected properties are part of the subject matter as a whole. *In re Papesch*, 315 F.2d 381, 137 USPQ 43 (CCPA 1963).

Analysis

The Claims

Claims 1-3, 5-35, 37, 38, 43-46 and 69-73 are set forth above. Claim 39 is directed to a combination that includes a fusion protein containing a nucleotide binding domain operatively linked to a ligand binding domain derived from an intracellular receptor where the nucleotide binding domain is a polydactyl zinc finger peptide or modular portion thereof that specifically interacts with a contiguous nucleotide sequence of at least about 3 nucleotides and the fusion protein is a ligand activated transcriptional regulator; or a nucleic acid molecule containing a sequence of nucleotides that encodes the fusion protein; and a regulatable expression cassette that contains at least one response element recognized by the nucleic acid binding domain of the fusion protein. Dependent claims 40-42 specify that the regulatable expression cassette contains a gene that encodes a therapeutic product; and that the fusion protein or nucleic acid encoding the fusion protein and the expression cassette are either in a single composition or in separate compositions.

Thus, all the claims contain as elements (1) a fusion protein containing a nucleotide binding domain operatively linked to a ligand binding domain; and (2) a fusion protein that is a ligand activated transcriptional regulator.

As discussed below, there is no teaching or suggestion in Beerli *et al.* for the construction nor the use of fusion proteins containing a ligand binding domain, whether derived from an intracellular receptor or otherwise. Further, contrary to the Examiner's assertion, there is no teaching or suggestion or

suggestion in Beerli *et al.* for regulating transcription by ligand activation, much less its application to any diagnostics or therapeutics, including gene therapy.

Teachings of the cited art and differences from the claimed subject matter

Beerli *et al.*

As discussed above, Beerli *et al.* is directed to the construction of polydactyl zinc finger fusion proteins that specifically bind to 9 or 18 bp target nucleotide sequences in a gene-specific manner. Beerli *et al.* teaches conversion of polydactyl zinc finger proteins into transcriptional activators or transcriptional repressors by fusion with KRAB, SID or ERD repressor domains, or VP16 or VP64 activation domains. Beerli *et al.* further teaches that the fusion proteins containing zinc finger domains fused to "effector" domains such as transcriptional activators or repressors function as "zinc finger-based transcriptional switches" (page 14629) for gene-specific transcriptional activation or repression. Beerli *et al.* also teaches that gene-specific transcriptional regulator fusion proteins provided therein can be applied to gene therapy, functional genomics and the generation of transgenic organisms.

Beerli *et al.* does not teach or suggest any fusion proteins, nucleic acid molecules encoding fusion proteins, vectors or combinations containing fusion proteins where the fusion proteins containing ligand binding domains. The ligand binding domains of the fusion proteins provided herein are derived from intracellular receptors such as, for example, the estrogen or progesterone receptors. Beerli *et al.* does not teach or suggest ligand activated transcriptional regulation.

There is no motivation, teaching or suggestion in Beerli *et al.* for the construction of fusion proteins that contain ligand binding domains and that are ligand activated transcriptional regulators or "gene switches," nor is there a suggestion in Beerli *et al.* to do that which Applicant has done.

It is alleged that Beerli *et al.* teaches a "general strategy" for the production of gene switches and that there would be motivation to use the nucleic acid molecules encoding the fusion proteins provided in Beerli *et al.* in gene therapy methods. The "general strategy" provided by Beerli *et al.*,

however, is that of gene-specific transcription that is controlled by zinc finger binding domains that are specific for particular gene sequences. Beerli *et al.* teaches that target genes, such as, for example, the *erbB-2* gene that is implicated in various cancers, may be transcriptionally activated or repressed as desired by targeting a transcriptional activator or transcriptional repressor domain to the target gene using a zinc finger domain fused to the transcriptional activator or repressor where the zinc finger domain recognizes the target gene in a sequence-specific manner.

There is no teaching or suggestion in Beerli *et al.* for the construction of fusion proteins that contain zinc finger domains operatively linked to ligand binding domains or operatively linked to both ligand binding domains and transcription regulation domains. Further, Beerli *et al.* does not teach or suggest any gene-specific transcriptional regulation where transcription is regulated by binding of a ligand to the ligand binding domain of a fusion protein. The ligand binding domains of the fusion proteins provided and claimed herein are derived from intracellular receptors such as, for example, the estrogen or progesterone receptors. The fusion proteins provided and claimed herein are directed to specific target genes by engineering their zinc finger nucleotide binding domains for target gene-specific sequence recognition, and transcription of the target gene is regulated by binding of a ligand to the ligand binding domain. The ligand binding domain of the fusion proteins provided and claimed herein may also be modified to alter its ligand specificity relative to the native hormone receptor. Beerli *et al.* does not teach or suggest any mechanism of gene-specific transcriptional regulation by ligand activation.

Thus, Beerli *et al.* fails to teach or suggest several elements of the claims, all of which (claims directed to fusion proteins, nucleic acid molecules encoding fusion proteins, vectors or combinations containing fusion proteins where the fusion proteins contain ligand binding domains) involve fusion proteins that contain a ligand binding domain and that are ligand activated transcriptional regulators. Beerli *et al.* provides no teaching or suggestion for modification of

its fusion proteins that would have resulted in the instantly claimed fusion proteins that include a ligand binding domain. Hence Beerli *et al.* does not teach or suggest a ligand activated transciptional regulator as claimed in the instant application, particularly one that is activated by a non-native ligand.

Further, because Beerli *et al.* does not teach or suggest fusion proteins containing ligand binding domains, nor is there any teaching or suggestion of ligand activated transcriptional regulation. There is no teaching or suggestion to use fusion proteins that contain ligand binding domains and that are ligand binding transcriptional regulators in diagnostic or therapeutic combinations or methods, including gene therapy. Therefore, since there is no motivation or suggestion in the cited reference to do that which Applicant has done (*In re Fritch*), the Examiner has failed to set forth a case of *prima facie* obviousness.

* * *

In view of the above amendments and remarks, reconsideration and allowance of the application are respectfully requested.

Respectfully submitted,
HELLER EHRMAN WHITE & McAULIFFE LLP

By:_____
Stephanie Seidman
Registration No. 33,779

Attorney Docket No.: 22908-1227B
Address all correspondence to:
HELLER EHRMAN WHITE & McAULIFFE LLP
4350 La Jolla Village Drive, 6th Floor
San Diego, CA 92122
Telephone: 858 450-8400
Facsimile: 858 587-5360
email:sseidman@HEWM.com

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Applicant: **Barbas III et al.**

Serial No.: 09/586,625

Filed: June 2, 2000

For: **LIGAND ACTIVATED TRANSCRIPTIONAL REGULATOR PROTEINS**

Art Unit: 1646

Examiner: Murphy, J.

CERTIFICATE OF MAILING BY "EXPRESS MAIL"
"Express Mail" Mailing Label Number
EV338077957US
Date of Deposit May 5, 2003

I hereby certify that this paper and the attached papers are being deposited with the United States Postal "Express Mail Post Office to Addressee" Service under 37 C.F.R. §1.10 on the date indicated above and addressed to:

Commissioner for Patents
U.S. Patent and Trademark Office
P.O. Box 1450
Alexandria, VA 22313-1450

Jayshree Aiyar Gerken

Jayshree Aiyar Gerken

ATTACHMENTS TO RESPONSE TO OFFICE ACTION

1. Marked-up paragraphs and claims (37 C.F.R. § 1.121)
2. Steitz, "*Structural Studies of Protein-Nucleic Acid Interactions*" (Cambridge University Press, 1990; pp 4-10)

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Applicant: **Barbas III et al.**

Serial No.: 09/586,625

Filed: June 2, 2000

For: **LIGAND ACTIVATED TRANSCRIPTIONAL REGULATOR PROTEINS**

Art Unit: 1646

Examiner: Murphy, J.

CERTIFICATE OF MAILING BY "EXPRESS MAIL"

"Express Mail" Mailing Label Number EV338077957US

Date of Deposit May 5, 2003

I hereby certify that this paper and the attached papers are being deposited with the United States Postal "Express Mail Post Office to Addressee" Service under 37 C.F.R. §1.10 on the date indicated above and addressed to:

Commissioner for Patents
U.S. Patent and Trademark Office
P.O. Box 1450
Alexandria, VA 22313-1450

05/05/03
Date

Jayshree Aiyar Gerken

Jayshree Aiyar Gerken

MARKED UP PARAGRAPHS AND CLAIMS (37 C.F.R. § 1.121)

IN THE SPECIFICATION:

Please amend the specification as follows:

Please amend the paragraph on page 1, lines 8-30, with the following:

Intracellular receptors are a superfamily of related proteins that mediate the nuclear effects of a variety of hormones and effector molecules, [include]including steroid hormones, thyroid hormones and vitamins A and D. Members of this family of intracellular receptors are prototypical ligand activated transcription factors. These receptors contain two primary functional domains: a DNA binding domain (DBD) that contains about sixty-six amino acids and a ligand-binding domain (LBD) located in the carboxyl-terminal half of the receptor that has about 300 amino [acids]acids. The receptors are inactive in the absence of hormone (ligand) by virtue of association with inactivating factors, such as heat shock proteins. Upon ligand binding, the receptors dissociate from the inactivating complex and dimerize, which renders them able to bind to DNA and modulate transcription.

Please amend the paragraph on page 2, lines 8-20, with the following:

Modified steroid hormone receptors have been developed for use for regulated expression of transgenes (see, *e.g.*, U.S. Patent No. 5,874,534 and published International PCT application No. WO 98/18925, which is based on

U.S. provisional application Serial No. 60/029,964) by modifying the ligand specificity of the LBD. In addition, the DNA binding domain of the receptor has been replaced with a non-mammalian DNA binding domain selected from yeast GAL4 DBD, a viral DBD and an insect DBD [binding domain] to provide for regulated expression of a co-administered gene containing a region recognized by the non-mammalian DBD. These constructs, however, have several drawbacks. The non-mammalian DBD is potentially immunogenic and the array of sequences recognized by these DBD is limited, thereby severely restricting gene targets.

Please amend the paragraph on page 2, lines 25-31, with the following:

Polypeptides that function as ligand activated transcriptional regulators and nucleic acid molecules encoding such polypeptides are provided. The polypeptides are fusion proteins that are ligand activated transcriptional [regulator]<u>regulators</u> that can be targeted to any desired endogenous or exogenous gene. Variants of the fusion protein can be designed to have different selectivity and sensitivity for endogenous and exogenous ligands.

Please amend the paragraph beginning on page 4, line 23, through page 5, line 4, with the following:

[Ligan]<u>Ligand</u> Binding Domain (LBD)

The LBD is derived from an intracellular receptor, particularly a steroid hormone receptor. The receptors from which the LBD is derived include, but is not limited to, glucocorticoid receptors, mineralocorticoid receptors, thyroid hormone receptors, retinoic acid receptors, retinoid X receptors, Vitamin D receptors, COUP-TF receptors, ecdysone receptors, Nurr-I receptors, orphan receptors and variants thereof. Receptors of these types include, but are not limited to, estrogen receptors, progesterone receptors, glucocorticoid-α receptors, glucocorticoid-β receptors, androgen receptors and thyroid hormone receptors. LBDs preferably are modified to alter ligand specificity so that they preferentially bind to an exogenous ligand, such as a drug, compared to an endogenous ligand.

Please amend th paragraph on page 5, lines 10-20, with the following:

The LBD is preferably modified so that it does not bind to the endogenous ligand for the receptor from which the LBD is derived, but to a selected ligand to permit fine tuned regulation of targeted genes. Hence, in certain embodiments, the ligand-binding domain has been modified to change its ligand selectivity compared to its [selective]selectivity in the native receptor. Preferably the modified ligand-binding domain is not substantially activated by endogenous ligands. Any method for altering ligand specificity, including systematic sequence alteration and testing for specificity, and selection protocols (see, *e.g.*, U.S. Patent No. 5,874,534 and Wang *et al.* (1994) *Proc. Natl. Acad. Sci. U.S.A. 91*:8180-8184) can be used.

Please amend the paragraph on page 6, lines 13-19 with the following:

Cys_2His_2 (C2H2) type zinc finger proteins are exemplary of the zinc fingers that can replace the naturally occurring DNA binding domain in an intracellular receptor, such as the C4-C4 type [domian]domain in a steroid receptor, to form a functional ligand-responsive transcription factor fusion protein. By virtue of the zinc finger, the resulting fusion protein exhibits altered DNA binding specificity compared to the unmodified intracellular receptor.

Please amend the paragraph on page 6, lines 24-30, with the following:

In preferred embodiments the zinc-finger portion of the fusion protein binds to a nucleotide sequence of the formula $(GNN)_n$, where G is guanidine, N is any nucleotide and n is an integer from 1 to 6, and typically n is 3 to 6. Preferably, the zinc-finger modular unit is derived from C2H2 zinc-finger peptide. More preferably, the zinc-finger peptide is a C2H2 zinc-finger peptide and has at least 90% sequence identity to a human zinc-finger peptide.

Please amend the paragraph on page 7, lines 9-19, with the following:

The transcription regulating domain can be any such domain known to [regulator]regulate or prepared to regulate eukaryotic transcription. Such TRDs are known, and include, but are not limited to, VP16, VP64, TA2, STAT-6, p65, and derivatives, multimers and combinations thereof that exhibit transcriptional regulation properties. The transcription regulating domain can be derived from

an intracellular receptor, such as a nuclear hormone receptor transcription activation (or repression) domain, and is preferably a steroid hormone receptor transcription activation domain or variant thereof that exhibits transcriptional regulation properties. Transcription domains include, but are not limited to, TAF-1, TAF-2, TAU-1, TAU-2, and variants thereof.

Please amend the paragraphs on page 8, lines 1-25, with the following:

Nucleic acid constructs

Also provided are nucleic acid molecules that encode the resulting fusion proteins. The nucleic acids can be included in vectors, suitable for expression of the proteins and/or vectors suitable for gene therapy. [Cell]Cells containing the vectors are also provided. Typically the cell is a eukaryotic cell. In other embodiments, the cell is a prokaryotic cell.

Also provided are expression cassettes that contain a gene of interest, particularly a gene encoding a therapeutic product, such as an angiogenesis inhibitor, operatively linked to a transcriptional regulatory region or response element, including sequences of nucleic acids to which a fusion [proteins]protein provided herein binds and controls transcription, particularly upon binding of a ligand to the LBD of the fusion polypeptide. Such expression cassettes can be included in a vector for gene therapy, and are intended for administration with, before or after, administration of the fusion protein or nucleic acid encoding the fusion protein. Genes of interest for exogenous delivery typically encode therapeutic proteins, such as growth factors, growth factor inhibitors or antagonists, tumor necrosis factor (TNF) inhibitors, anti-tumor agents, angiogenesis agents, anti-angiogenesis agents, clotting factors, apoptotic and other suicide genes.

Compositions, combinations and kits

Also provided are compositions that contain the fusion proteins or the vectors that [encoded]encode the fusion proteins. Combinations of the fusion proteins or nucleic acids encoding the proteins and nucleic acid encoding a targeted gene with regulatory regions selected for activation by the fusion protein are also provided.

Pl ase amend the paragraph beginning on pag 9, lin 26, through page 10, line 9, with the following:

Methods

Methods for regulating expression of endogenous and exogenous genes are provided. The methods are practiced by administering to a cell a composition that contains an effective amount or concentration of the fusion protein or of nucleic acid molecule, such as a vector that encodes the fusion protein. The nucleic acid binding domain (DBD) of the fusion protein is selected to bind to a targeted nucleic acid sequence in the genome of the cell or in an exogenously administered nucleic acid molecule, and the transcription regulating domain (TRD) is selected to regulate transcription from a selected promoter, which typically is operatively linked to the targeted nucleic acid binding domain. The exogenously administered nucleic acid molecule comprises an expression cassette encoding a gene of interest and operatively linked to a regulatory region that contains elements, such as a promoter and response elements.

Please amend the paragraph on page 10, lines 16-21, with the following:

At the same time or at a later time, a composition [containing] comprising a ligand that binds to the ligand binding domain of the fusion protein is also administered. The ligand can be administered in the same composition as the fusion protein (or encoding nucleic acid molecule) or in a separate composition. The ligand and fusion protein may be administered sequentially, simultaneously or intermittently.

Please amend the paragraphs on page 11, lines 1-29, with the following:

In other embodiments, the methods for regulating gene expression in a cell are effected by administering to the cell a composition containing an effective amount of the nucleic acid molecule that encodes the ligand activated transcriptional regulatory fusion protein, a regulatable expression cassette containing a gene operatively linked to at least one response element for the gene recognized by the nucleotide binding domain of the polypeptide encoded by the polynucleotide, and a pharmaceutically acceptable excipient; and

administering to the cell a ligand that binds to the ligand binding domain of the encoded polypeptide, where the nucleotide binding domain of the encoded polypeptide [to] binds to the response element and activates or represses transcription of the gene.

Methods for treating a cellular proliferative disorder by the *ex vivo* introduction of a recombinant expression vector encoding the fusion protein are provided. Cellular proliferative [disorder]disorders include disorders associated with transcription of a gene at reduced or increased levels.

Administration [can] of the composition(s) can be effected *in vitro*, *in vivo* or *ex vivo*. One such method includes the removal of a tissue sample from a subject with a disorder, such as a cell proliferative disorder, isolating hematopoietic or other cells from the tissue sample, and contacting isolated cells with the fusion protein or a nucleic acid molecule encoding the fusion protein, and, optionally, a target specific gene. Optionally, the cells can be treated with a growth factor, such as interleukin-2 for example, to stimulate cell growth, before reintroducing the cells into the subject. When reintroduced, the cells specifically target the cell population from which they were originally isolated. In this way, the trans-repressing activity of the zinc finger-nucleotide binding polypeptide may be used to inhibit or suppress undesirable cell proliferation in a subject. Preferably, the subject is a human.

Please amend the paragraph on page 15, lines 3-16, with the following:

As used herein, the ligand binding domain (LBD) of the fusion proteins provided herein refers to the portion of the fusion protein responsible for binding to a selected ligand. The LBD optionally and preferably includes dimerization and inactivation functions. The LBDs in the proteins herein are derived from the 300 amino acid carboxyl-terminal half of intracellular receptors, particularly those that are members of the steroid hormone nuclear receptor superfamily. It is the portion of the receptor protein with which a ligand interacts thereby inducing a cascade of events leading to the specific association of an activated receptor with regulatory elements of target genes. In these receptors the LDB includes the hormone binding function, the inactivation [funciton]function, such

as through interactions with heat shock proteins (hsp), and dimerization function. The LBDs used herein include such LBDs and modified derivatives thereof, particularly forms with altered ligand specificity.

Please amend the paragraphs beginning on page 15, line 24, through page 16, line 17, with the following:

As used herein, the DNA binding domain (DBD), or alternatively the nucleic acid (or nucleotide) binding domain, refers to the portion of the fusion polypeptide provided herein that provides specific nucleic acid binding capability. The use of the abbreviation DBD is not meant to limit it to DNA binding domains, but is also [intented]intended to include polypeptides that bind to RNA. The nucleic acid binding domain functions to target the protein to specific genes by virtue of the specificity of the interaction of the TRD region for nucleotide sequences operatively linked to the transcriptional apparatus of a gene. The DBD targets the fusion protein to the selected targeted gene or genes, which gene(s) may be endogenous or exogenously added.

As used herein, operatively linked means that elements of the fusion polypeptide, for example, are linked such that each perform or [functios]function as intended. For example, the repressor is attached to the binding domain in such a manner that, when bound to a target nucleotide via that binding domain, the repressor acts to inhibit or prevent transcription. Linkage between and among elements may be direct or indirect, such as via a linker. The elements are not necessarily adjacent. Hence a repressor domain of a TRD can be linked to a DNA binding domain using any linking procedure well known in the art. It may be necessary to include a linker moiety between the two domains. Such a linker moiety is typically a short sequence of amino acid residues that provides spacing between the domains. So long as the linker does not interfere with any of the functions of the binding or repressor domains, any sequence can be used.

Please amend the paragraphs beginning on page 17, line 12, through page 18, line 8, with the following:

In a peptide or protein, suitable conservative substitutions of amino acids are known to those of skill in this art and may be made generally without

altering the biological activity of the resulting molecule. Those of skill in this art recognize that, in general, single amino acid substitutions in non-essential regions of a polypeptide do not substantially alter biological activity (see, *e.g.*, Watson *et al. Molecular Biology of the Gene*, 4th Edition, 1987, The [Bejacmin]Benjamin/Cummings Pub. co., p.224).

As used herein, a delivery plasmid is a plasmid vector that carries or delivers [nucleotide]nucleic acids encoding a therapeutic gene or gene that encodes a therapeutic product or a precursor thereof or a regulatory gene or other factor that results in a therapeutic effect when [delived]delivered *in vivo* in or into a cell line, such as, but not limited to a packaging cell line, to propagate therapeutic viral vectors.

As used herein, "recombinant expression vector" or "expression vector" refers to a plasmid, virus or other vehicle known in the art that has been manipulated by insertion or incorporation of heterologous DNA, such as nucleic acid encoding the fusion proteins herein or expression cassettes provided herein. Such expression vectors contain a [promotor]promoter sequence for efficient transcription of the inserted nucleic acid in a cell. The expression vector typically contains an origin of replication, a promoter, as well as specific genes that permit phenotypic selection of transformed cells.

As used herein, a DNA or nucleic acid homolog refers to a [a] nucleic acid that includes a preselected conserved nucleotide sequence, such as a sequence encoding a therapeutic polypeptide. By the term "substantially homologous" is meant having at least 80%, preferably at least 90%, most preferably at least 95% homology therewith or a less percentage of homology or identity and conserved biological activity or function.

Please amend the paragraphs on page 19, lines 1-30, with the following:

Whether any two nucleic acid molecules have nucleotide sequences that are at least 80%, 85%, 90%, 95%, 96%, 97%, 98% or 99% "identical" can be determined using known computer algorithms such as the "FAST A" program, using for example, the default parameters as in Pearson *et al.* (1988) *Proc. Natl. Acad. Sci. USA 85*:2444. Alternatively the BLAST function of the

National Center for Biotechnology Information database may be used to determine [identity]identify.

In general, sequences are aligned so that the highest order match is obtained. "Identity" *per se* has an art-recognized meaning and can be calculated using published [techniques. (See]techniques (see, *e.g.*: *Computational Molecular Biology*, Lesk, A.M., ed., Oxford University Press, New York, 1988; *Biocomputing: Informatics and Genome Projects*, Smith, D.W., ed., Academic Press, New York, 1993; *Computer Analysis of Sequence Data, Part I*, Griffin, A.M., and Griffin, H.G., eds., Humana Press, New Jersey, 1994; *Sequence Analysis in Molecular Biology*, von Heinje, G., Academic Press, 1987; and *Sequence Analysis Primer*, Gribskov, M. and Devereux, J., eds., M Stockton Press, New York, 1991). While there exist a number of methods to measure identity between two polynucleotide or polypeptide sequences, the term "identity" is well known to skilled artisans (Carillo *et al.* (1988) *SIAM J Applied Math 48*:1073). Methods commonly employed to determine identity or similarity between two sequences include, but are not limited to, those disclosed in Guide to Huge Computers, Martin J. Bishop, ed., Academic Press, San Diego, 1994, and Carillo *et al.* (1988) *SIAM J Applied Math 48*:1073. Methods to determine identity and similarity are codified in computer programs. Preferred computer program methods to determine identity and similarity between two sequences include, but are not limited to, GCG program package (Devereux, J., *et al., Nucleic Acids Research 12(I)*:387 (1984)), BLASTP, BLASTN, and FASTA (Atschul, S.F., *et al., J Molec Biol 215*:403 (1990)).

Please amend the paragraphs beginning on page 20, line 28, through page 22, line 2, with the following:

As used herein, genetic therapy involves the transfer of heterologous DNA to the certain cells, target cells, of a mammal, [particulaly]particularly a human, with a disorder or conditions for which such therapy is sought. The DNA is introduced into the selected target cells in a manner such that the heterologous DNA is expressed and a therapeutic product encoded thereby is produced. Alternatively, the heterologous DNA may in some manner mediate

expression of DNA that encodes the therapeutic product, or it may encode a product, such as a peptide or RNA that in some manner mediates, directly or indirectly, expression of a therapeutic product. Genetic therapy may also be used to deliver nucleic acid encoding a gene product that replaces a defective gene or supplements a gene product produced by the mammal or the cell in which it is introduced. The introduced nucleic acid may encode a therapeutic compound, such as a growth factor inhibitor thereof, or a tumor necrosis factor or inhibitor [thereor]thereof, such as a receptor therefor, that is not normally produced in the mammalian host or that is not produced in therapeutically effective amounts or at a therapeutically useful time. The heterologous DNA encoding the therapeutic product may be modified prior to introduction into the cells of the afflicted host in order to enhance or otherwise alter the product or expression thereof. Genetic therapy may also involve delivery of an inhibitor or repressor or other modulator of gene expression.

As used herein, heterologous DNA is DNA that encodes RNA and proteins that are not normally produced *in vivo* by the cell in which it is expressed or that mediates or encodes mediators that alter expression of endogenous DNA by affecting transcription, translation, or other regulatable biochemical processes. Heterologous DNA may also be referred to as foreign DNA. Any DNA that one of skill in the art would recognize or consider as heterologous or foreign to the cell in which it is expressed is herein encompassed by heterologous DNA. Examples of heterologous DNA include, but are not limited to, DNA that encodes traceable marker proteins, such as a protein that confers drug resistance, DNA that encodes therapeutically effective substances, such as anti-cancer agents, enzymes and hormones, and DNA that encodes other types of proteins, such as antibodies. Antibodies that are encoded by heterologous DNA may be secreted or expressed on the surface of the cell in which the heterologous DNA has been introduced.

Please amend the paragraph on page 23, lines 1-16, with the following:

As used herein, isolated with reference to a nucleic acid molecule or polypeptide or other biomolecule means [thatthe]that the nucleic acid or

polypeptide has separated from the genetic environment from which the polypeptide or nucleic acid were obtained. It may also mean altered from the natural state. For example, a polynucleotide or a polypeptide naturally present in a living animal is not "isolated," but the same polynucleotide or polypeptide separated from the coexisting materials of its natural state is "isolated", as the term is employed herein. Thus, a polypeptide or polynucleotide produced and/or contained within a recombinant host cell is considered isolated. Also intended as an "isolated polypeptide" or an "isolated polynucleotide" are polypeptides or polynucleotides that have been purified, partially or substantially, from a recombinant host cell or from a native source. For example, a recombinantly produced version of a [compounds]compound can be substantially purified by the one-step method described in Smith *et al.* (1988) *Gene 67:*31-40. The terms isolated and purified are sometimes used interchangeably.

Please amend the paragraph beginning on page 23, line 24, through page 24, line 2, with the following:

Isolated or purified as it refers to preparations made from biological cells or hosts means any cell extract containing the indicated DNA or protein including a crude extract of the DNA or protein of interest. For example, in the case of a protein, a purified preparation can be obtained following an individual technique or a series of preparative or biochemical techniques and the DNA or protein of interest can be present at various degrees of purity in these preparations. The procedures may include for example, but are not limited to, ammonium sulfate fractionation, gel filtration, ion exchange [change] chromatography, affinity chromatography, density gradient centrifugation and electrophoresis.

Please amend the paragraph on page 28, lines 14-23, with the following:

As used herein with regard to nucleic acid molecules, including DNA fragments, the phrase "operatively linked" means the sequences or segments have been covalently joined, preferably by conventional phosphodiester bonds, into one strand of DNA, whether in single or double stranded form such that operatively linked portions [functions]function as intended. The choice of vector

to which a transcription unit or a cassette provided herein is operatively linked depends directly, as is well known in the art, on the functional properties desired, *e.g.*, vector replication and protein expression, and the host cell to be transformed, these being limitations inherent in the art of constructing recombinant DNA molecules.

Please amend the paragraph on page 29, lines 26-30, with the following:

As used herein, ligand refers to any compound that interacts with the ligand binding domain of a receptor and [modulate]modulates its activity; ligands typically activate receptors. Ligand can also include compounds that activate the receptor without binding. A natural ligand is a compound that normally interacts with the receptor.

Please amend the paragraph on page 30, lines 4-11, with the following:

As used herein, non-natural ligands or non-native ligands refer to compounds that are normally [are] not found in mammals, such as humans, that bind to or interact with the ligand binding domain of a receptor. Hence, the term "non-native ligands" refers to those ligands that are not naturally found in the specific organism (man or animal) in which gene therapy is contemplated. For example, certain insect hormones such as ecdysone are not found in humans. As such ecdysone is non-native hormone to an animal, such as a human.

Please amend the paragraph on page 31, lines 1-4, with the following:

As used herein, administration of a therapeutic composition can be effected by any means, and includes, but is not limited to, subcutaneous, intravenous, intramuscular, intrasternal, infusion techniques, [intraperitoneally]intraperitoneal administration and parenteral administration.

Please amend the paragraphs beginning on page 31, line 22, through page 32, line 11, with the following:

The fusion protein also includes a LBD that is derived from an intracellular receptor, preferably a hormone receptor, more preferably a steroid receptor. The LBD can be modified to have altered ligand specificity so that endogenous or natural ligands do not interact with it, but non-natural ligands do. The fusion

protein also can include a transcription regulating domain (TRD) that regulates transcription of the targeted gene(s). In some embodiments, the TRD can repress transcription of an endogenous gene; in others it can activate expression of an endogenous or exogenous gene.

Hence the fusion protein is made by operably linking a LBD domain from an [intacellular]intracellular receptor to [a] one or more zinc finger domains, selected to bind to a targeted gene. A transcription regulating domain can also be operably linked. This is accomplished by any method known to those of skill in the art. Generally the fusion protein is produced by expressing nucleic acid encoding the fusion protein.

1. **Ligand Binding Domain (LBD)**

The ligand binding domain is derived from an intracellular receptor, and is preferably derived from a nuclear hormone receptor. The LBD of an intracellular receptor includes the approximately 300 amino acids from the carboxy [terminal]terminus, which can be used with or without modification.

Please amend the paragraphs beginning on page 32, line 28, through page 33, line 25, with the following:

The LBD can be modified by deletion of from about 1 up to about 150, typically 120, amino acids on the carboxyl terminal end of the receptor from which the LBD derives. Systematic deletion of amino acids and [subsuqent]subsequent testing of the ligand specificity and of the resulting LBD can be used to empirically identify mutations that lead to modified LBDs that have desired properties, such as preferential interaction with non-natural ligands. Exemplary mutations are described in the Examples herein, and also are known to those of skill in the art (see, *e.g.*, U.S. Patent No. 5,874,534; U.S. Patent No. 5,935,934; U.S. Patent No. 5,364,791; and International PCT application No. 98/18925, which is based on U.S. provisional application Serial No. 60/029,964; International PCT application No. 96/40911, which is based on U.S. application Serial No. 08/479,913l) and references cited therein. Hence a LBD or modified form thereof prepared by known methods is obtained and operably linked to a DBD; a TRD is also linked as needed.

2. Nucleic Acid Binding Domain (DBD)

Zinc fingers are modular nucleic acid binding peptides. The zinc fingers, or modules thereof, or [variant]variants thereof can be used to construct fusion proteins that specifically interact with targeted sequences. Zinc fingers are ubiquitous proteins, and many are well-characterized. For example, methods and rules for preparation and selection of zinc fingers based upon the C2H2 class of zinc fingers with unique specificity are known (see, *e.g.*, International PCT application No. WO 98/54311 and International PCT application No. 95/19431; see, also U.S. Patent No. 5,789,538; Beerli *et al.* (1999) *Proc. Natl. Acad. Sci. U.S.A. 96*:2758-2763; Beerli *et al.* (1995) *Proc. Natl. Acad. Sci. U.S.A. 95*:14628-14633; see, also U.S. application Serial No. 09/173,941, filed 16 October, 1998, published as International PCT application No. WO 00/23464). Exemplary targeting sequences are provided herein.

Please amend the paragraph beginning on page 34, line 16, through page 35, line 2, with the following:

For example, zinc finger variants can be prepared by identifying a zinc finger or modular unit thereof, creating an expression library, such as a phage display library (see, *e.g.*, International PCT application No. WO 98/54311, Barbas *et al.* (1991) *Methods 2*:119; Barbas *et al.* (1992) *Proc. Natl. Acad. Sci. U.S.A. 89*:4457), encoding polypeptide variants of the zinc finger or modular [unit therof]units thereof, expressing the library in a host and screening for variant peptides having a desired specificity. Zinc fingers may also be constructed by combining amino acids (or encoding nucleic acids) according to the known rules of binding specificity and, if necessary, testing or screening the resulting peptides to ensure the peptide has a desired specificity. Because of the modular nature of zinc fingers, where each module can be prepared to bind to a three nucleotide [squence]sequence, peptides of any specificity can be prepared from the modules. The number of modules used depends upon the specificity of gene targeting desired. Modular units are combined; spacers (*i.e.* TGEKP, TGQKP) required to maintain spacing and conformational features of the modular domains are included in the peptide (see, *e.g.*, WO 98/54311).

Please amend the paragraph on page 36, lines 5-13, with the following:

A zinc finger-nucleotide binding peptide domain contains a unique heptamer (contiguous sequence of 7 amino acid residues) within the α-helical domain of the polypeptide, which heptameric sequence determines binding specificity to a target [necleotide]nucleotide. The heptameric sequence can be located anywhere within the α-helical domain but it is preferred that the heptamer extend from position -1 to position 6 as the residues are conventionally numbered in the art. A peptide nucleotide-binding domain can include any β-sheet and framework sequences known in the art to function as part of a zinc finger protein.

Please amend the paragraph on page 40, lines 15-21, with the following:

More often, however, two or three amino acids are selected for nucleotide recognition. His3 or Lys3 (and to a lesser extent, Gly3) are selected for the recognition of a middle guanine. Ser3 and Ala3 are selected to recognize a middle thymine. Thr3, Asp3, and Glu3 are selected to recognize a middle cytosine. Asp and Glu [were] are selected in position -1 to recognize a 3′ cytosine, while Thr-1 and Ser-1 are selected to recognize a 3′ thymine.

Please amend the paragraph on page 41, lines 13-23, with the following:

Further the data demonstrate that sequence motifs at positions -1,1, and 2 rather than the simple identity of the position 1 residue are required for highly specific recognition of the 3′ base. These residues likely provide the proper stereo-chemical context for interactions of the helix in terms of recognition of specific bases and in the exclusion of other bases, the net result being highly specific interactions. Ready recombination of the disclosed domains then allows for the creation of proteins, typically [polypdactyl]polydactyl proteins, of defined specificity precluding the need to develop phage display libraries in their generation. Such family of zinc finger domains is sufficient for the construction of 16 or 17 million proteins that bind to the 5′-$(GNN)_6$-3′ family of DNA sequences.

Please amend the paragraphs beginning on page 45, line 16, through page 46, line 19, with the following:

e. Screening of [varint]variant zinc finger and other DBD peptides

Any method known to those of skill in the art for identification of functional modular domains derived from zinc fingers and combinations thereof can be employed. An exemplary method for identifying variants of zinc fingers or other polypeptides that bind to zinc finger binding motifs is provided. Components used in the method include a nucleic acid molecule encoding a putative or modified zinc finger peptide operably linked to a first inducible promoter and a reporter gene operably linked to a second inducible promoter and a zinc finger-nucleotide binding motif, wherein the incubating is carried out under conditions sufficient to allow the components to interact, and measuring the affect of the putative DBD peptide on the expression of the reporter gene is provided.

For [exampole]example, a first inducible promoter, such as the arabinose promoter, is operably linked to the nucleotide sequence encoding the putative DBD polypeptide. A second inducible promoter, such as the lactose promoter, is operably linked to a zinc finger derived-DNA binding motif followed by a reporter gene, such as β-galactosidase. Incubation of the components may be *in vitro* or *in vivo*. *In vivo* incubation may include prokaryotic or eukaryotic systems, such as *E.coli* or COS cells, respectively. Conditions that allow the assay to proceed include incubation in the presence of a substance, such as arabinose and lactose, which activate the first and second inducible promoters, respectively, thereby allowing expression of the nucleotide sequence encoding the putative trans-modulating protein nucleotide sequence. Determination of whether the putative modulating protein binds to the zinc finger-nucleotide binding motif, which is operably linked to the second inducible promoter, and affects its activity is measured by the expression of the reporter gene. For example, if the reporter gene is β-galactosidase, the presence of blue or white plaques indicates whether the putative modulating protein enhances or inhibits,

respectively, gene expression from the promoter. Other commonly used assays to assess the function from a promoter, including chloramphenicol acetyl transferase (CAT) assay, are known to those of skill in the art. Prokaryote and eukaryote systems can be used.

Please amend the paragraph beginning on page 50, line 20, through page 51, line 2, with the following:

Transcriptional repressors are well known in the art, and any such repressor can be used herein. The repressor is a polypeptide that is operatively linked to the nucleic acid binding domain as set forth above. The repressor [in]is operatively linked [ot]to the binding domain in that it is attached to the binding domain in such a manner that, when bound to a target nucleotide via that binding domain, the repressor acts to inhibit or prevent transcription. The repressor domain can be linked to the binding domain using any linking procedure well known in the art. It may be necessary to include a linker moiety between the two domains. Such a linker moiety is typically a short sequence of amino acid residues that provides spacing between the domains. So long as the linker does not interfere with any of the functions of the binding or repressor domains, any sequence can be used.

Please amend the paragraphs on page 52, lines 1-23, with the following:

c. Activators

Exemplary and preferred transcription activation domains include any protein or factor that regulates transcription. [Examplary]Exemplary transcriptional regulation domains include, but are not limited to, VP16, TA2, VP64, STAT6 [amd]and relA.

4. Exemplary construct based on human integrin β3 and erbB-2 target sequences

To exemplify the generation of zinc finger modular [dmomains]domains and peptides containing one or more of such domains to produce peptides with DNA binding specificity and therapeutic potential, target sequences have been identified based on human integrin β3 and *erb*B-2 (Ishii *et al.* (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*:4374-4378) genomic sequences.

Integrin β3 as a target for cancer gene therapy

Integrin $\alpha_v\beta_3$ is the most [promiscous]promiscuous member of the integrin family and has been identified as a marker of angiogenic vascular tissue. For instance, integrin $\alpha_v\beta_3$ shows enhanced expression on blood vessels in human wound granulation tissue but not in normal skin. Following the induction of angiogenesis, blood vessels show a four-fold increase in $\alpha_v\beta_3$ expression compared to blood vessels not undergoing this process. It has been reported that a cyclic peptide or monoclonal antibody antagonist of integrin $\alpha_v\beta_3$ blocks cytokine- or tumor-induced angiogenesis on the chick chorioallantoic membrane. Therefore, inhibition of integrin $\alpha_v\beta_3$ expression provides an approach to block tumor-induced angiogenesis.

Please amend the paragraph beginning on page 55, line 22, through page 56, line 4, with the following:

In another embodiment, gene therapy can be accomplished using a combination of the vectors described above. For example, a retroviral vector can deliver a stably integrated, inducible transgene cassette into a population of cells either *in vitro* (*ex vivo*) or *in vivo*. Subsequently, the integrated transgene can be activated by transducing this same cell population with a second vector, such as an adenovirus vector capable of expressing the fusion protein, followed by the administration of the specific ligand inducing agent. This is [is] particularly useful where "one time" activation of the transgene is desired, for example as a cellular suicide mechanism. An example of this application is the stable integration of an inducible transgene cassette containing the herpes simplex virus thymidine kinase gene (HSV Tk). Subsequent activation of this gene confers sensitivity to ganciclovir and allows ablation of this modified cell.

Please amend the paragraphs beginning on page 57, line 20, through page 58, line 12, with the following:

Since recombinant retroviruses are defective, they require assistance in order to produce infectious vector particles. This assistance can be provided, for example, by using helper cell lines that contain plasmids encoding all of the structural genes of the retrovirus under the control of regulatory sequences

within the LTR. These plasmids are missing a nucleotide sequence which enables the packaging mechanism to recognize an RNA transcript for [encapsitation]encapsidation. Helper cell lines which have deletions of the packaging signal include but are not limited to Ψ2, PA317 and PA12, for example. These cell lines produce empty virions, since no genome is packaged. If a retroviral vector is introduced into such cells in which the packaging signal is intact, but the structural genes are replaced by other genes of interest, the vector can be packaged and vector virion produced. The vector virions produced by this method can then be used to infect a tissue cell line, such as NIH 3T3 cells, to produce large quantities of chimeric retroviral virions.

b. Nonviral Delivery systems

"Non-viral" delivery techniques for gene therapy include DNA-ligand complexes, adenovirus-ligand-DNA complexes, direct injection of DNA, $CaPO_4$ precipitation, gene gun techniques, electroporation, liposomes and lipofection. Any of these methods are available to one skilled in the art and would be suitable for use herein. Other suitable methods are available to one skilled in the art, and it is to be understood that the methods herein may be accomplished using any of the available methods of transfection.

Please amend the paragraph on page 60, lines 14-18, with the following:

In general, the compounds bound to the surface of the targeted delivery system are ligands and receptors [perimitting]permitting the targeted delivery system to find and "home in" on the desired cells. A ligand may be any compound of interest that interacts with another compound, such as a receptor.

Please amend the paragraph on page 61, lines 15-25, with the following:

Administration of [a] nucleic acid molecules provided herein to a target cell *in vivo* may be accomplished using any of a variety of techniques well known to those skilled in the art. The vectors of the methods herein may be administered orally, parentally, by inhalation spray, rectally, or topically in dosage unit formulations containing conventional pharmaceutically acceptable carriers, adjuvants, and vehicles. Suppositories for rectal administration of the drug can be prepared by mixing the drug with a suitable non-irritating excipient

such as cocoa butter and polyethylene glycols that are solid at ordinary temperatures but liquid at the rectal temperature and therefore melt in the rectum and release the drug.

Please amend the paragraphs beginning on page 62, line 27, through page 63, line 27, with the following:

3. Ligands

As noted, the ligands may be naturally-occurring ligands, but are preferentially non-natural ligands with which the LBD is modified to [specificallly]specifically interact. Methods for modifying the LBD are known, as are methods for screening for such ligands.

Ligands include, non-natural ligands, hormones, anti-hormones, synthetic hormones, and other such compounds. Examples of non-natural ligands, anti-hormones and non-native ligands include, but are not limited to, the following: 11β-4-dimethylaminophenyl)-17α-hydroxy-17α-propinyl-4,9-e stradiene-3-one (RU38486 or Mifepestone); 11β-(4-dimethylaminophenyl)-17α-hydroxy-17β-(3-hydroxypropyl)-13α-methyl-4,9-gonadiene-3-one (ZK98299 or Onapristone); 11β-(4-acetylphenyl)-17β-hydroxy-17α-(1-propinyl)-4,9-estr adiene-3-one (ZK112993); 11β-(4-dimethylaminophenyl)-17β-hydroxy-17α-(3-hydroxy-1 (Z)-propenyl-estra-4,9-diene-3-one (ZK98734); (7β11β,17β)-11-(4-dimethylaminophenyl)-7-methyl-4',5'-dihydrospiroy 'ester-4,9-diene-17,2' (3'H)-furan!-3-one (Org31806); (11β,14β,17α)-4',5'-dihydro-11-(4-dimethylamino-phenyl)y'spi roestra-4,9-diene-17,2'(3'H)-furan!-3-one (Org31376); 5-alpha-pregnane-3,2-dione. Additional non-natural ligands include, in general, synthetic non-steroidal estrogenic or anti-estrogenic compounds, broadly defined as selective estrogen receptor modulators (SERMS). Exemplary [coumpounds]compounds include, but are not limited to, tamoxifen and raloxifen.

4. Pharmaceutical compositions and combinations

Also provided is a pharmaceutical composition containing a therapeutically effective amount of the fusion protein, or a nucleic acid molecule encoding the fusion protein in a pharmaceutically acceptable carrier.

Pharmaceutical compositions containing one or more fusion proteins with different zinc finger-nucleotide binding domains are [contemplatd]contemplated. Also provided are pharmaceutical compositions containing the expression cassettes, and also compositions containing the ligands. Combinations containing a plurality of compositions are also provided.

Please amend the paragraph on page 65, lines 5-15, with the following:

D. Methods of gene regulation

[Method]Methods of regulating expression of endogenous and exogenous genes are provided. In particular, ligand-dependent methods are provided. In practicing the methods, a target [nucleotide]nucleic acid molecule containing a sequence that interacts with the nucleic acid binding domain of the fusion protein exposed to an effective amount of the fusion protein in the presence of an effective binding amount of a ligand, which can be added simultaneous with or subsequent to the fusion protein. The nucleic acid binding domain of the fusion protein binds to a portion of the target nucleic acid [moleucule]molecule and the ligand binds to the ligand binding domain of the fusion protein. Exposure can occur *in vitro, in situ* or *in vivo*.

Please amend the paragraph beginning on page 66, line 11, through page 67, line 2, with the following:

Treatments

Methods for gene therapy are provided. The fusion proteins are administered either as a protein or as a nucleic acid encoding the protein and delivered to cells or tissues in a mammal, such as a human. The fusion protein is targeted either to a specific sequence in the genome (an endogenous gene) or to an exogenously added gene, which is [administerd]administered as part of an expression cassette. Prior to, simultaneous with or subsequent to [adminstration]administration of the fusion protein, a ligand that specifically interacts with the LBD in the fusion protein is [adminstered]administered. In embodiments, in which the targeted gene is exogenous, the expression cassette, which can be present in a vector, is administered, simultaneous with or subsequent to [adminstration]administration of the fusion protein. These

methods are intended for treatment of any genetic disease, for treatment of acquired disease and any other conditions. Diseases include, cell proliferative disorders, such as cancer. Such therapy achieves its therapeutic effect by introduction of the fusion protein that includes the zinc finger-nucleotide binding polypeptide, either as the fusion [or] protein or encoded by a nucleic acid molecule that is expressed in the cells, into cells of animals having the disorder. Delivery of the fusion protein or nucleic acid molecule can be effected by any method known to those of skill in the art, including methods described herein. For example, it can be effected using a recombinant expression vector such as a chimeric virus or a colloidal dispersion system.

Please amend the paragraph beginning on page 74, line 17, through page 75, line 2, with the following:

Selection strategy for the generation of six-finger proteins with DNA binding specificity

Based on the modular nature of zinc finger domains, as well as the fact that each zinc finger recognizes 3 bp of DNA sequence, several strategies can be employed to generate zinc finger proteins, with preferably one to three fingers, with desired DNA binding specificity[an]. For instance, *in vitro* evolution of a six-finger protein binding an 18bp target sequence can follow the strategy outlined in FIGURE 1. The target sequence is divided into six 3bp sub-sites, A-F. In the first step, a Zif268-based zinc finger phage display library in which the central finger 2 is randomized is selected against all 6 subsites in the context of the 2 wild type fingers. After [successfull]<u>successful</u> generation of all the finger 2 variants required for a given target, cDNAs encoding three-finger proteins recognizing either half-site 1 (ABC) or half-site 2 (DEF) are constructed via PCR overlap extension. Finally, standard cloning procedures are used to construct a gene encoding a six-finger protein recognizing the whole 18bp target site.

Please amend the paragraph on page 84, lines 1-15, with the following:

EXAMPLE 2

Construction Of Fusion Proteins Containing Zinc Finger Domains and Transcriptional Repressors And Activators

In order to [demonsrate]demonstrate use of zinc finger proteins as gene-specific transcriptional regulators, the E2C(Sp1), B3B(Sp1), and B3C2(Sp1) six-finger proteins were fused to a number of effector domains (Beerli *et al.* (1998) *Proc. Natl. Acad. Sci. U.S.A.95*:14628-14633). Transcriptional repressors were generated by attaching either of three human-derived repressor domains to the zinc finger protein. The first repressor protein was prepared using the ERF repressor domain (ERD) (Sgouras *et al.* (1995) *EMBO J. 14*:4781-4793), defined by amino acids 473 to 530 of the *ets2* repressor factor (ERF). This domain mediates the antagonistic effect of ERF on the activity of transcription factors of the *ets* family. A synthetic repressor was constructed by fusion of this domain to the C-terminus of the zinc finger protein.

Please amend the paragraph beginning on page 85, line 7, through page 86, line 6, with the following:

Specific regulation of erbB-2 promoter activity

Reporter constructs containing fragments of the *erbB-2* promoter coupled to a luciferase reporter gene were generated to test the specific activities of the *erbB-2* specific synthetic transcriptional regulators. The target reporter plasmid contained nucleotides -758 to -1 with respect to the ATG initiation codon, whereas the control reporter plasmid contained nucleotides -1571 to -24, thus lacking all but one nucleotide of the E2C binding site encompassed in positions -24 to -7. Both promoter fragments displayed similar activities when transfected transiently into HeLa cells, in agreement with previous observations. To test the effect of zinc finger-repressor domain fusion constructs on *erbB-2* promoter activity, HeLa cells were transiently co-transfected with each of the zinc finger expression vectors and the luciferase reporter constructs (Beerli *et al.*, (1998) *Proc. Natl. Acad. Sci. U.S.A.95*:14628-14633). Significant repression was observed with each construct. The ERD and SID fusion proteins produced approximately 50% and 80% repression, respectively. The most potent repressor was the KRAB fusion protein. This protein caused complete repression of *erbB-2* promoter activity. The observed residual activity was at the background level of the promoter-less pGL3 reporter. In contrast, none of the

proteins caused significant repression of the control *erbB-2* reporter construct lacking the E2C target site, demonstrating that repression is indeed mediated by specific binding of the E2C(Sp1) protein to its target site. Expression of a zinc finger protein lacking any effector domain resulted in weak repression, approximately 30%, indicating that most of the repression observed with the SID and KRAB constructs is caused by their effector domains, rather than by DNA-binding alone. This observation strongly suggests that the mechanism of repression is active inhibition of transcription initiation rather than of elongation. Once initiation of transcription by RNA polymerase II has [occured]occurred, the zinc finger protein appears to be readily displaced from the DNA by the action of the polymerase.

Please amend the paragraph on page 86, lines 17-25, with the following:

Based on the efficient and specific regulation of a reporter construct driven by the *erbB-2* promoter, the effect of transiently transfected zinc finger expression plasmids on activity of the [endogeneous]endogenous *erbB-2* promoter was analyzed. As a read-out of *erbB-2* promoter activity, ErbB-2 protein levels were analyzed by Western blotting. Significantly, E2C(Sp1)-VP64 lead to an upregulation of ErbB-2 protein levels, while E2C(Sp1)-SKD lead to its downregulation. This regulation was specific, since no effect was observed on expression of EGFR.

Please amend the paragraph on page 87, lines 5-17, with the following:

Specific regulation of integrin β3 promoter activity

To test the activity of transcriptional regulators specific for the integrin β3 promoter, a reporter plasmid was constructed containing the luciferase open reading frame under control of the integrin β3 promoter. When compared to the two *erbB-2* promoter fragments described above, the integrin β3 promoter fragment had a very low activity. In fact, in some experiments no activation of luciferase expression over background was detected, preventing an analysis of the effects of the KRAB fusion proteins. However, when the VP64 fusion proteins were tested an efficient activation of the integrin β3 promoter was observed. B3B(Sp1)-VP64 and B3C2(Sp1)-VP64 stimulated transcription 12 and

[22fold]22-fold, respectively. Activation of transcription was specific, since no effect on the activity of the *erbB-2* promoter was detected.

Please amend the paragraph on page 88, lines 3-9, with the following:

The original PR DNA binding domain can be replaced by engineered zinc finger proteins. For example, the three finger protein Zif268(C7) was fused to the N-terminus of the PR ligand binding domain (PBD) (aa 640 to 914), and the VP16 activation domain to its C-terminus. It was found that this fusion protein [protein] was able to regulate an SV40 promoter luciferase construct with ten upstream Zif268(C7) binding sites in an RU486-dependent manner.

Please amend the paragraph beginning on page 89, line 30, through page 90, line 10, with the following:

In order to optimize the ability of the fusion proteins to regulate gene expression, it may be necessary to add additional heterologous transactivating domains to the receptor. To facilitate these studies, fusion proteins were constructed either with the full length LBD extending to estrogen receptor residue 595, or with LBD fragments truncated at amino acid (aa) 554 to remove the F region. The full-length constructs are referred to as long (L), the truncated versions as short (S). All constructs contain a heterologous transactivation domain (TA) comprised of a VP16 minimal domain, unless otherwise noted, fused to the carboxy terminus of the ligand binding domain. VP16 minimal domain trimer has the amino acid residue sequence 3 x (PADALDDFDLDML) (SEQ ID NO: [36]47), and is the tetracycline controlled transactivator (tTA) TA2 (Baron *et al.* (1997) *Nucleic Acids Research 25*:2723-2729).

Please amend the paragraph beginning on page 91, line 10, through page 92, line 3, with the following:

The general cloning strategy was as follows. Three fragments (A, B, and C with reference to FIG. 3) of human estrogen receptor ligand binding domain (LBD) with or without the F region were built into the pcDNA3.1 (Invitrogen) vector backbone through a series of PCR amplification and cloning steps. Initially the LBD fragment A without F region (i.e. short form; LBDAS) and with F region (i.e. long form; LBDAL) were PCR amplified from a plasmid clone of the

human wild type estrogen receptor, pHEGO (Tora *et al. EMBO J. 8*:1981-1986) with primer pairs NR1/NR2 and NR1/NR3 respectively (Table 1). Convenient restriction sites were incorporated into primers (Table 1) as needed. The PCR amplified LBDAS and LBDAL fragments were first cloned into the *Srf* I site of pCR-ScriptAmpSK(+) vector ([Strategene]Stratagene), resulting in constructs pLBDAS and pLBDAL. The VP16 minimal domain trimer (TA2; Baron *et al.* (1997) *Nucleic Acids Research 25*:2723-2729) was PCR amplified from plasmid pTTA2 (Clontech) with primer pairs NR4 and NR9 and cloned into the *Spl*I and *Not*I site of pLBDAS and pLBDAL to generate pLBDASTA2 and pLBDALTA2. To generate LBD fragment B without the F region (LBDBS) and LBD fragment C without the F region (LBDCS), PCR primers NR7 and NR8, which represent the 5' boundary of the LBD region fragment in chimerics B and C respectively were designed (Table 6, below). These primers were paired with the 3' end primer NR6, which incorporates a unique *Blp*I site in ER. PCR fragments from pHEGO with primer pair NR6/NR7 and PCR fragment with NR6/NR8 were then cloned into the *Spe*I and *Blp*I site of pLBDC7ASTA2 backbone. This resulted in plasmid pLBDBSTA2 and pLBDCSTA2.

Please amend the paragraphs beginning on page 94, line 3, through page 95, line 16, with the following:

A DNA oligonucleotide containing two inverted binding sites for the C2H2 domain known as C7, each half site separated by 3 bp, was used for the initial assessment of DNA binding. This palindromic configuration mimics the composition of the native estrogen receptor response element (ERE), except that the natural 6 bp half site of ERE is replaced by the 9 bp half site specified by C7. Binding of the C7-LBD fusion proteins A, B, and C, all in the short form, were tested and compared to the control proteins C7VP16 and 2C7VP16 (see, Liu, *et al.* (1997) *Proc. Natl. Acad. Sci. U.S.A 94*:5525-5530, which describes the control proteins). For each protein, binding was tested in the absence or presence of 100 fold excess of unlabeled oligonucleotide (1.75 μM) as a competitor. Competition of the gel shift product by the unlabeled oligonucleotide indicates the band is a specific protein:DNA interaction. The

results demonstrated that C7VP16 can bind once or twice to the oligonucleotide, creating two specific gel shift bands. 2C7VP16 binds only once to the oligonucleotide containing two inverted C7 sites. Notably, C7LBDA and C7LBDB bind strongly to yield one major species, which runs higher than any of the control bands. Although true molecular mass cannot be determined from this type of mobility assay, the relative size of the complexes suggest the protein bound for C7LBD is larger than for C7VP or 2C7VP. The size of the band and presence of only one major species indicate that the fusion [protien]protein ZFP-LBD is binding to the oligonucleotide as a dimer. No significant gel shift product was detected for C7LBD chimeric C, suggesting that the addition of the additional native zinc finger from the estrogen receptor may have reduced the affinity of the fusion protein for its C2H2-specific DNA binding site. Finally, the reduction of binding for each of the gel shift products by the addition of the unlabeled oligonucleotide indicates that these fusion proteins are binding to DNA in a sequence specific manner.

To further demonstrate that the chimera ZFP-LBD binds to DNA as a dimer, the binding of C7LBD A, B, and C to oligonucleotides containing one or two C7 binding sites was tested. Three fusion proteins (C7LBDAS, C7LBDBS and C7LBDCS) were tested against three different target oligonucleotide sequences, which contained one C7 half site or two C7 half sites either in palindromic or direct repeat orientation.

Oligo 1: gat cca aag tcg cgt ggg cgc agc gcc cac gcg atc aaa ga (SEQ ID NO: 48)

Oligo 2: gat cca aag tcc agg cga gcg cgt ggg cgg cag atc aaa ga (SEQ ID NO: 49)

Oligo 3:gat cca aag tcg cgt ggg cgc agg cgc gag cgt ggg cgg atc aaa ga (SEQ ID NO: 50) Insert hard return

Gel shift assay conditions were the same as the standard [procotol]protocol described above. The results showed that C7LBDAS and C7LBDBS were able to bind to both oligonucleotides containing two C7 half

sites, but not to the oligo containing only one half site. C7LBDCS bound weakly or not at all to all three targets.

Please amend the paragraph on page 98, lines 15-25, with the following:

The ability of C7LBD short form chimeric proteins A, B, and C to regulate reporter gene expression in an estrogen-dependent manner was studied in Cos and HeLa cells. The constitutive activators C7VP16 and 2C7VP16 were used as positive controls. The results show that the three ZFP-LBD fusion proteins gave a similar profile in Cos and [Hela]HeLa cells. All three ZFP-LBD fusion proteins had an estrogen dependent effect on the luciferase reporter gene. The characteristic pattern is that A has greater total activity than B and B has greater total activity than C. Likewise, the basal or ligand-independent effect of these proteins on the reporter gene follows a similar pattern[;]: A>B>C. The estrogen dependent effect on gene expression ranged from two-fold to nine-fold in these experiments.

Please amend the paragraphs beginning on page 99, line 28, through page 100, line 8, with the following:

TabAssays were performed with [Hela]HeLa cells transfected with 0.5 ug of 6x2C7pGL3Luc reporter and 0.1 μg regulator, Luc activity was determined as previously described. When the human STAT6 transactivation domain was used to replace the TA2 VP minimal domain trimer, the same low basal activity and 9 fold ligand dependent induction of transgene, two-fold less than with the TA2 domain, was obtained.

The [incorpration]incorporation of NLS upstream of the full length VP16 (FIG. 24, C7ASnlsVP16) greatly increased the folding induction compared to TA2 or VP16 without the NLS, but the total activity was significantly decreased. When the full length VP16 domain was used, it gave about 2 fold higher total activity, but high basal activity resulting in weaker ligand dependent induction (3-fold).

Please amend the paragraph on page 109, lines 16-24, with the following:

All four fusion constructs were fully sequenced and tested in a luciferase assay for their ability to regulate the *erbB-2* promoter in a ligand-dependent manner. It was found that the PR 6 Finger heterodimer was inactive; a similar observation was made with an C7-RxR // EcR-VP16 heterodimer. In contrast, the E2C-ER // ER-VP64 heterodimer had some activity, and the addition of Tamoxifen lead to a roughly three-fold upregulation of promoter activity. Variations in the ratio of the two heterodimerization partners led to an increased [inducability]inducibility, up to total of 5.3-fold.

Please amend the paragraphs beginning on page 110, line 26, through page 112, line 12, with the following:

For all transfections, HeLa cells were plated in 24-well dishes and used at a confluency of 40-60%. Typically, 175 ng reporter plasmid (pGL3-[promotor]promoter constructs or, as negative control, pGL3basic) and 25 ng effector plasmid (zinc finger constructs in pcDNA3 or, as negative control, empty pcDNA3.1) were transfected using the Lipofectamine reagent (Gibco BRL). Cell extracts were prepared approximately 48 hours after transfection. Luciferase activity was measured with the Promega luciferase assay reagent in a MicroLumat LB96P luminometer (EG&G Berthold).

Bombyx mori EcR

A plasmid (LNCVBE) containing the coding region for *Bombyx mori* EcR was obtained from F. Gage. *Bombyx mori* EcR is PCR amplified from this plasmid using the primers listed below and AmpliTaq DNA Polymerase (Hoffmann-LaRoche). Forward and backward primers were chosen to allow construction of the constructs corresponding to FIG. 14 but replacing [*Drosohila*]*Drosophila* EcR by *Bombyx mori* EcR.

(FseI)-BE: (SEQ ID NO: 63)
GAGGAGGAGGGCCGGCCGGAGGCCTGAATGTGTCATACAGGAGCCC

(SfiI)-BE: (SEQ ID NO: 64)
GAGGAGGAGGGCCCAGGCGGCCAGGCCTGAATGTGTCATACAGGAGCCC

BE-(AscI): (SEQ ID NO: 65)

GAGGAGGAGGGCGCGCCCCTCCGCCACGTCCCAGATCTCCTCGAG

C7-R-VP16 // C7-E-VP16

This [hetereodimer]heterodimer was examined on two reporters, one containing 10 C7 sites and one containing 6 2C7 sites, and in two cell lines, HeLa and NIH. In all cases the C7-R-VP16 construct alone showed a high activation of transcription (840-fold) that did not depend on the presence of Ponasterone A. However the C7-E-VP16 construct showed a very little activation of transcription on its own. C7-R-VP16 // C7-E-VP16 together showed the same behavior as C7-R-VP16 alone.

C7-R // E-VP16

In this [hetereodimer]heterodimer, the activation domain on RXR is dropped to eliminate the basal activation observed above. EcR has no DNA-binding domain to render activation dependent on the presence of DNA-bound RXR. This [hetereodimer]heterodimer was tested with the 3-finger protein C7 on the 10C7 reporter and with the 6-finger protein E2C on the E2P reporter that contains a single E2P binding site. In both cases no significant activation could be observed.

C7-R // C7-E-VP16

To combine the low basal activity of C7-R // E-VP16 with the high activation seen with C7-R-VP16 // C7-E-VP16, the activation domain on RXR was dropped but the zinc finger protein on EcR was retained. In this set-up, on a 6x2C7 reporter, a 5-fold activation with very low basal activity was observed. Similar constructs using the more powerful VP64 activation domain have also been made.

E2C- ER // ER-VP64

This heterodimeric [onstruct]construct showed 5.3 fold tamoxifen-dependent activation at ratios of 6.7/60 and 2.2/60 of the erbB-2 promoter.

Please amend the paragraph on page 119, lines 5-20, with the following:

A series of different reporter constructs assembled in order to determine the optimal target DNA spacing and orientation of the C2H2 binding sites for

transgene induction, C7LBDBS was transfected into HeLa cells and assayed for basal and tamoxifen induced activity on a series of reporter constructs [diagramed]diagrammed above. Reporter constructs were constructed by cloning double stranded oligonucleotides containing the various binding sites into the multiple cloning site of the pGL3Luc reporter. "Response elements" composed of direct, inverted (palindromic), and everted repeats of two C7 binding sites were compared; each response element was separated by two (2) bp except in the control 6 X 2C7, where spacing was 5 bp. Several arrays of directly repeated single C7 sites were tested with various spacing. The data show that direct repeats and everted repeats are preferred over palindromic binding sites. Further, 6 C7 sites, each separated by 2 bp is comparable to the control element of 6 x 2C7, even though it contains only half the number of individual C7 binding sites.

Please amend the paragraph on page 120, lines 8-23, with the following:

Protein binding to DNA was analyzed by gel shift assay. The electrophoretic studies used 2C7 recombinant molecular constructs using native PAGE and SDS PAGE analysis of binding to a DNA probe [containg]containing six 2C7 binding sites. In this experiment, the 2C7VP16 protein was used as a control and the P32-labeled DNA probe was the 6x2C7 fragment excised from the 6X2C7pGL3Luc. Sufficient 2C7VP protein was added to yield three distinct gel shifted products. When a similar level of protein for the 2C7LBD A, B, and C were applied, only a single weak band was observed. By comparison to the one and two copies bound bands for the 2C7VP16 control, the 2C7LBD band position suggests it is binding as a monomer. Furthermore, the weak level of binding compared to the 2C7VP16 control suggests the DNA binding affinity of the 2C7 domain is significantly reduced in the context of the LBD fusion protein. Results of *in vitro* expressed proteins by SDS-PAGE, indicated equal amounts of proteins expressed and the expected relative increase in size for the LBD A, B, and C forms.

Please amend the paragraph beginning on page 122, line 29, through page 123, line 10, with the following:

Left end shuttle plasmid construction for ZFP-LBD Fusion Protein Regulators

Shuttle plasmids containing the left viral ITR, CMV immediate [ealy]early promoter and ZFP-LBD regulator were prepared in the plasmid pAvCVIx (Figure 26). Note that this vector contains a loxP recombination site just downstream of the poly adenylation sequence. DNA encoding the intact reading frame for the chimeric regulators C7LBD As(G521R), C7LBD Bs(G521R), and C7LBD Bs(G400V) were excised from the appropriate pCDNA constructions, (see figures 4 and 5 for LBD As and LBD Bs constructs respectively) by digestion with restriction enzymes EcoRI and Not I. The ZFP-LBD DNA fragments were modified with Klenow to fill in the restriction site overhangs and blunt end ligated into the EcoRV at bp 1393 site of pAvCvIx to generate pAvCv-C7LBD As(G521R), pAvCv-C7LBD Bs(G521R), and pAvCv-C7LBD Bs(G400V).

Please amend the paragraph on page 125, lines 4-23, with the following:

In Vitro Regulation with Adenovirus Vectors

The ability to regulate expression of a transgene delivered by an adenovirus vector was demonstrated by the following experiment. [Hela]HeLa cells were infected with a mixture of two adenovirus vectors, one containing a fusion protein regulator either (Av3-C7LBD-A(G521R) or Av3-C7LBD-B(G52R), the other containing the 6x2C7SV40-luc cassette. To determine the optimal ratio of target vector to effector vector, two different doses of the transgene or target vector (50 or 250 viral particles per cell) at three different ratios of effector vector (50, 250, 750 particles per cell for each target dose) were tested. Twenty four hours after vector transduction, the cells were treated where appropriate with 100 nM 4-OH-tamoxifen. Following an additional 24 hrs incubation, the cells were lysed and assayed for luciferase activity. For the Av3CV-C7LBD A(G521R) vector, the data indicate relatively low levels of luc expression in the absence of 4-OHT, a strong 4-OHT dependent induction and a dose dependent increase in luc activity as more fusion protein regulator vector is used. At the highest doses (750 particles per cell) of chimeric regulator vector

tested, tamoxifen-specific induction of 460 to 560 fold over basal was achieved at target vector doses of 250 and 50 particles per cell, respectively.

Please amend the paragraph on page 126, lines 15-23, with the following:

Details of the animal study are as [follow]follows. On study day 1, C57Bl/6 male mice were given a total adenovirus vector dose of 2 x 10^{11} particles via tail vein injection. On day two blood samples were collected, then animals were injected i.p. with 200 [ul]μl of sunflower seed oil containing 5% DMSO and either no, 50 [ug]μg, or 500 [ug]μg of tamoxifen (Sigma # T56448). Blood samples were collected daily for three days following drug administration, and on study days 8 and 10. At the completion of the study, murine endostatin levels were determined by ELISA (Accucyte Kit, Cytimmune Sciences, Maryland).

Please amend the paragraph beginning on page 127, line 4, through page 128, line 5, with the following:

In addition, groups 5 and 6 were similar to groups 3 and 4, but animals received 0.5 x 10^{11} of the Av3TATA-mEndo vector and 1.5 x 10^{11} of the C7LBD regulator vector, for a 1:3 ratio of target to effector. Groups 3 – 6 each contained no drug, 50 [ug]μg, and 500 [ug]μg tamoxifen treatment sub-groups.

The results showed a dramatic induction of murine endostatin following the day 2 administration of 50 μg of tamoxifen. The highest level of induction was observed on day 3, the day immediately following drug administration. Compared to the basal level observed on day 3 in the no tamoxifen groups, the C7LBDAs(G521R) and C7LBDBs(G400V) regulators gave comparable fold induction, approximately 17 fold, and comparable absolute levels of expression, around 1500 ng/ml. In this study, the endogenous murine endostatin levels in an untreated mouse cohort was 20 $\pm$7 ng/ml. The drug-induced endostatin expression rapidly declines by day 5, three days after drug administration, which is presumably due to the clearance of the tamoxifen and biological half life of the endostatin protein. In contrast, expression in the Av3RSV-mEndo treatment group persists at 200 ng/ml through day 15. In the 1:3 target to

effector ratio groups, tamoxifen-induced expression reached 600 – 900 ng/ml, approximately 1/2 the level in the 1:1 ratio cohorts. This result indicates that *in vivo*, the transgene-containing vector, not the fusion protein-encoding vector, is limiting for absolute protein expression. Furthermore, endostatin expression in the animals treated with 500 μg tamoxifen was comparable to the animals treated with only 50 μg, indicating that the lower dose of tamoxifen is sufficient to fully activate the As(G521R) and Bs(G400V) regulators. Finally, the comparable low basal level of endostatin observed in the As(G521R) and Bs(G400V) groups suggests that the endogenous level of estrogen in the C57Bl/6 mice is not sufficient to induce the [estogen]estrogen-responsive Bs(G400V) regulator. An elevation in basal endostatin levels observed at days 3 – 5 appeared to be a non-specific effect resulting from adenovirus vector administration, since the Av3Null vector has an effect similar to the Av3TATA-mEndo containing groups.

Please amend the paragraph beginning on page 128, line 18, through page 129, line 2, with the following:

EXAMPLE 20

Construction and evaluation of the Cys_2-His_2 Zinc finger DBD-ERLBD regulators in Lentiviral Vectors

In order to demonstrate controlled gene expression in an integrated vector system, the [the] regulatory system described in Example 19 with the adenoviral vectors were used to develop a series of lentiviral vectors. These vectors contained either the ZFP-LBD fusion protein linked to the immediate early CMV promoter or a regulatable transgene (either eGFP or luciferase) linked to the 6 X 2C7 array of C7 binding sites and either the minimal promoter from SV40 or C-fos TATA. The fusion protein-encoding vector and the regulatable transgene vector can then be used to generate lentiviral vector supernatant. The supernatant can be used to stably transduced human cells either singly or in parallel. Stable cell lines containing the integrated vectors can then be induced with the appropriate activating drug (*e.g.*, 4-OH-tamoxifen) and gene expression is measured as fold induction in the presence and absence of drug.

Please amend the paragraph beginning on page 130, line 19, through page 131, line 2, with the following:

Evaluation of the ZFP-LBD fusion proteins and regulatable lentiviral vectors

Transduction of HeLa cells by inducible lentiviral vectors

Subconfluent HeLa cells were transduced with either HIV6X2C7SvLuc or HIV6X2C7TATALuc vector supernatant for 24 hours followed by [tranduction]transduction with HIVAS521R lentiviral vector supernatant. Cells were allowed to recover from infection for 24 hours in fresh culture medium after which 4-OH-tamoxifen (100 or 1000 nm) was added to the culture for an additional 24 hours. Cells were lysed in a standard luciferase lysis buffer, subjected to freeze thaw and analyzed for luciferase activity using a luciferase assay kit (Promega). The results showed that cells infected with either HIV6X2C7SvLuc or HIV6X2C7TATALuc followed by transduction with HIVCMVAS521R resulted in a 13.1 and 11.7 fold stimulation in luciferase activity respectively, when given 4-OH-tamoxifen.

IN THE ABSTRACT:

Please amend the abstract as follows:

Please amend the paragraph on page 139, lines 1-9, with the following:

ABSTRACT

Fusion proteins for use as ligand-dependent transcriptional regulators are provided. The fusion proteins include a nucleotide binding domain operatively linked to a ligand-binding domain. They also can include a transcription regulating domain. The nucleotide binding domain is a zinc-finger peptide that binds to a targeted contiguous nucleotide sequence of from 3 to about 18 nucleotides are provided. The fusion proteins are used for gene therapy. Also provided are polynucleotides encoding the fusion proteins, expression vectors, and transfected cells.

IN THE CLAIMS:

Please amend claims 1, 8, 13, 20, 22, 25, 32, 39, 43, 45, 69 and 73 as follows:

1. (Amended twice) A fusion protein, comprising a nucleotide binding domain operatively linked to a ligand binding domain [derived] from an intracellular receptor, wherein:

the nucleotide binding domain is a polydactyl zinc-finger peptide or modular portion thereof that [specifically] interacts with a contiguous nucleotide sequence of at least about 3 nucleotides;

the ligand binding domain has been modified to change its ligand specificity compared to the native hormone receptor; and

the fusion protein is a ligand activated transcriptional regulator.

8. (Amended three times) A fusion protein, comprising a nucleotide binding domain operatively linked to a ligand binding domain [derived] from an intracellular receptor, wherein:

the nucleotide binding domain is a polydactyl zinc-finger peptide or modular portion thereof that [specifically] interacts with a contiguous nucleotide sequence of at least about 3 nucleotides;

the zinc-finger peptide is comprised of modular units from a C2H2 zinc-finger peptide that [specifically] interacts with a sequence of nucleotides and targets the fusion protein to an exogenous or endogenous gene that comprises the sequence of nucleotides; and

the fusion protein is a ligand activated transcriptional regulator.

13. (Amended twice) The fusion protein of claim 3, wherein the hormone receptor is a progesterone receptor variant or an estrogen receptor variant, wherein a receptor variant comprises a ligand binding domain that has altered ligand specificity [selectivity and sensitivity] for endogenous and exogenous ligands [that differ from] <u>relative to</u> its native [ligands] <u>receptor</u>.

20. (Amended twice) A fusion protein, comprising a nucleotide binding domain operatively linked to a transcription regulating domain and a ligand binding domain [derived] from an intracellular receptor, wherein:

the nucleotide binding domain is a polydactyl zinc-finger peptide or modular portion thereof that [specifically] interacts with a contiguous nucleotide sequence of at least about 3 nucleotides;

the transcription regulating domain comprises a transcription repression domain; and

the fusion protein is a ligand activated transcriptional regulator.

22. (Amended twice) A fusion protein, comprising a nucleotide binding domain operatively linked to a transcription regulating domain and a ligand binding domain [derived] from an intracellular receptor, wherein

the nucleotide binding domain is a polydactyl zinc-finger peptide or modular portion thereof that [specifically] interacts with a contiguous nucleotide sequence of at least about 3 nucleotides;

the fusion protein is a ligand activated transcriptional regulator; and the fusion protein is encoded by the sequence of nucleotides set forth in any of SEQ ID Nos. 1-18.

25. (Amended three times) A nucleic acid molecule, comprising a sequence of nucleotides encoding a fusion protein, wherein:

the fusion protein comprises a nucleotide binding domain operatively linked to a ligand binding domain [derived] from an intracellular receptor, wherein the nucleotide binding domain is a polydactyl zinc-finger peptide or modular portion thereof that [specifically] interacts with a contiguous nucleotide sequence of at least about 3 nucleotides;

the fusion protein is a ligand activated transcriptional regulator; and

the fusion protein is encoded by a sequence of nucleotides set forth in [any of] SEQ ID [Nos. 1-18] <u>No. 1</u>.

32. (Amended twice) A viral vector comprising a sequence of nucleotides encoding a fusion protein, wherein:

the fusion protein comprises a nucleotide binding domain operatively linked to a ligand binding domain [derived] from an intracellular receptor, wherein the nucleotide binding domain is a polydactyl zinc-finger peptide or

modular portion thereof that [specifically] interacts with a contiguous nucleotide sequence of at least about 3 nucleotides; and

the fusion protein is a ligand activated transcriptional regulator.

39. (Amended twice) A combination, comprising:

a fusion protein comprising a nucleotide binding domain operatively linked to a ligand binding domain [derived] from an intracellular receptor, wherein

the nucleotide binding domain is a polydactyl zinc-finger peptide or modular portion thereof that [specifically] interacts with a contiguous nucleotide sequence of at least about 3 nucleotides and the fusion protein is a ligand activated transcriptional regulator; or

a nucleic acid molecule comprising a sequence of nucleotides that encodes the fusion protein; and

a regulatable expression cassette that comprises at least one response element recognized by the nucleic acid binding domain of the fusion protein.

43. (Amended twice) A composition for regulating gene expression, comprising an effective amount of:

a fusion protein comprising a nucleotide binding domain operatively linked to a ligand binding domain [derived] from an intracellular receptor, wherein the nucleotide binding domain is a polydactyl zinc-finger peptide or modular portion thereof that [specifically] interacts with a contiguous nucleotide sequence of at least about 3 nucleotides and the fusion protein is a ligand activated transcriptional regulator; or

a nucleic acid molecule comprising a sequence of nucleotides that encodes the fusion protein; and

a pharmaceutically acceptable excipient.

45. (Amended twice) A composition for regulating gene expression comprising an effective amount of:

a fusion protein comprising a nucleotide binding domain operatively linked to a transcription regulating domain and a ligand binding domain [derived] from

an intracellular receptor, wherein the nucleotide binding domain is a polydactyl zinc-finger peptide or modular portion thereof that [specifically] interacts with a contiguous nucleotide sequence of at least about 3 nucleotides and the fusion protein is a ligand activated transcriptional regulator; and

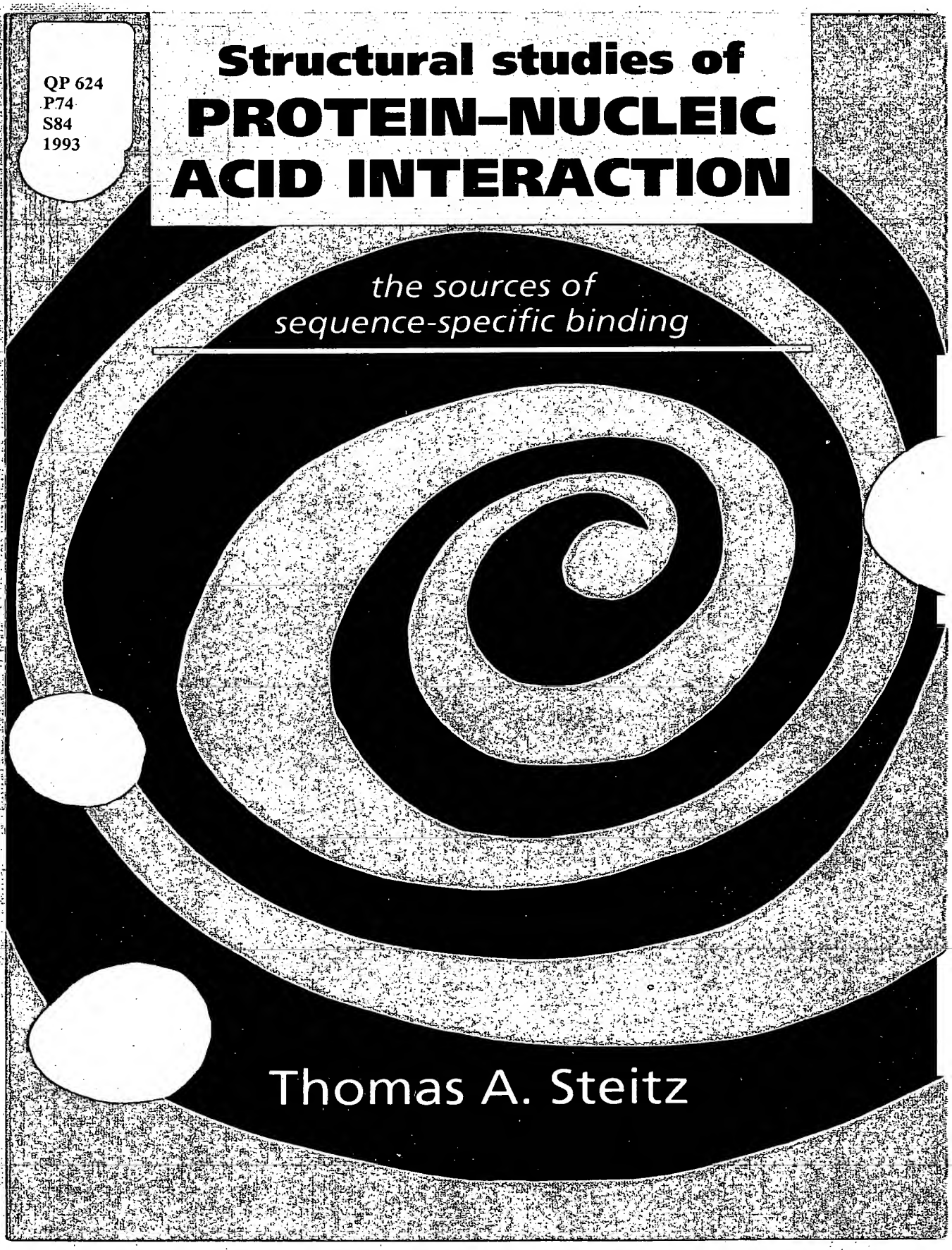
a pharmaceutically acceptable excipient.

69. (Amended) The fusion protein of claim 1, wherein the polydactyl zinc-finger peptide or modular portion thereof [specifically] interacts with a contiguous nucleotide sequence of at least about 3 nucleotides to about 18 nucleotides.

73. (Amended) The fusion protein of claim 9, wherein the zinc finger peptide comprised of at least one zinc finger or a variant thereof [specifically] binds to a targeted nucleic acid molecule with a dissociation constant of less than about 1.0 nanomolar.

Structural studies of PROTEIN–NUCLEIC ACID INTERACTION

the sources of sequence-specific binding



Thomas A. Steitz

Structural studies of protein–nucleic acid interaction

the sources of sequence-specific binding

THOMAS A. STEITZ

Department of Molecular Biophysics and Biochemistry
and Howard Hughes Medical Institute at Yale University

CAMBRIDGE
UNIVERSITY PRESS

Published by the Press Syndicate of the University of Cambridge
The Pitt Building, Trumpington Street, Cambridge CB2 1RP
40 West 20th Street, New York, NY 10011–4211, USA
10 Stamford Road, Oakleigh, Melbourne 3166, Australia

First published 1990 in *Quarterly Reviews of Biophysics*
First published as a book 1993

Printed in Great Britain at the University Press, Cambridge

A catalogue for this book is available from the British Library

Library of Congress cataloguing in publication data

Steitz, Thomas A.
Structural studies of protein–nucleic acid interaction: the sources of sequence-specific binding / Thomas A. Steitz.
p. cm.
Originally published in Quarterly reviews of biophysics in 1990.
Includes bibliographical references.
ISBN 0-521-41489-X (pbk.)
1. DNA-protein interactions. 2. RNA-protein interactions.
I. Title.
QP624.75.P74S84 1993
574.87'328 – dc20 92-41805 CIP

ISBN 0 521 41489 X paperback

UP

studies emanating from molecular genetic approaches to correlate structure with function.

Less rapid progress has been made on determining the structures of protein complexes with folded RNA molecules and less can be surmised about the nature of protein–RNA interaction from the structures of RNA-binding proteins alone. The structures of *B. stearothermophilus* tyrosyl-tRNA synthetase (tyrRS) (Bhat *et al.* 1982; Brick *et al.* 1989), *E. coli* methionyl-tRNA synthetase (MetRS) (Zelwer *et al.* 1982; Brunie *et al.* 1987) and *E. coli* elongation factor TU (Jurnak, 1985; La Cour *et al.* 1985) have been determined in the absence of tRNA. Recently, however, the structure of *E. coli* glutaminyl-tRNA synthetase (GlnRS) complexed to $tRNA^{Gln}$ and ATP has been determined and provides a wealth of information on the details of protein–RNA interactions (Rould *et al.* 1989; Perona *et al.* 1989).

Major contributors to this remarkable set of scientific advances have been technical advances in molecular genetics, DNA synthesis, increased speed and power of computation, rapid methods of data collection in protein crystallography and the development of 2D NMR methods of structure determination.

Clearly the ability to clone and overexpress the protein and RNA gene products for these normally rare nucleic acid-binding proteins has made accessible a large variety of proteins in this class. Co-crystallization of proteins with the ligands and substrates to which they bind has been a standard technique for examining the source of their specificity in detail over since the determination of the structure of lysozyme complexed with a trisaccharide was established in 1965 (Johnson & Phillips, 1965). Only after the technology of DNA synthesis advanced to the stage that tens of milligrams of a specific-sequence oligonucleotide could be made and purified in the early 1980s was it possible to co-crystallize proteins with DNA and carry out similar studies on DNA-binding proteins. The current availability of DNA-synthesizing machines and high-pressure liquid chromatography makes this technology now accessible to non-organic chemical laboratories. Similarly, the ability now to make large quantities of specific RNA species either by cloning its gene into a high expression vector or by *in vitro* transcription by T7 RNA polymerase makes the study of protein–RNA complexes accessible to crystallographic analysis. The advent of rapid X-ray crystallographic data collection techniques provided by 2D area detectors has only begun to have a dramatic effect on the pace of structural studies of macromolecules.

2. PRINCIPLES OF SEQUENCE-SPECIFIC NUCLEIC-ACID RECOGNITION

Structural, biochemical and molecular-genetic studies of protein–nucleic acid complexes have established two important sources of sequence specificity in protein–nucleic acid interactions: (1) Direct hydrogen bonding and van der Waals interaction between protein side chains and the exposed edges of base pairs, primarily in the major groove of B-DNA and to a lesser extent the minor grooves of DNA and RNA provides structural complementarity to correct but not to incorrect sequences. (2) The sequence-dependent bendability or deformability of

duplex DNA or RNA provides sequence selectivity by virtue of the ability of some nucleic-acid sequences to take up a particular structure required for binding to a protein at lower free energy cost than other sequences.

2.1 *The problem that is set: what is being recognized*

We are concerned here primarily with the interaction between proteins and duplex DNA and RNA, often in a nucleic-acid sequence-dependent fashion. The 3D structure of double-stranded DNA is highly polymorphic (Kennard & Hunter, 1989) but variants of two forms, A-form and B-form, are of relevance to the proteins described here. Fig. 1 shows an important difference between A- and B-form DNA. In B-DNA the major groove is wide enough to accommodate an α-helix and the functional groups on the exposed edges of the base pairs can be contacted by protein; the minor groove, on the other hand, is deep and narrow (5·7 Å) and thus less accessible to secondary structures such as an α-helix. For A-DNA (or RNA which is always A-form) the opposite is true. The minor groove is shallow and broad (10–11 Å wide for RNA) whereas the major groove is very deep and narrow [4 Å for RNA (Delarue & Moras, 1989)]. The width of the minor groove in B-DNA varies depending on base composition. AT rich sequences have a narrower minor groove (3·5 Å) than GC rich sequences (Yoon *et al.* 1988). In general, however, one might expect proteins to directly decode DNA sequences via interactions in the major groove but discriminate among RNA sequences via interactions in the minor groove. This appears for the most part to be the case.

A second important consideration in the suitability of the major and minor grooves for direct sequence recognition is the degree of structural variation of the four base pairs as viewed from the two grooves. Seeman *et al.* (1976) pointed out that the base pairs present a more richly varied set of hydrogen bond donors to the major groove as compared to the minor groove. Figure 2 shows that the minor groove side of the base pairs is a veritable recognition desert with only the N2 of guanine distinguishing AT from GC. The patterns of donors and acceptors on the major groove side, however, can distinguish all four base pairs. It is possible, therefore, that direct minor groove recognition can distinguish only a binary code (GC or CG *vs.* AT or TA) whereas major groove recognition can discriminate among all four base pairs.

2.2 *Role of the major groove in DNA recognition*

The extensive hydrogen bonding and shape complementarity between the major groove of B-DNA and the surfaces of many of the sequence-specific DNA-binding proteins has been extensively documented from high-resolution crystal structures of DNA complexes with both the 434 and λ repressor DNA-binding fragments (Aggarwal *et al.* 1988; Jordan & Pabo, 1988), the 434 *cro* repressor (Wolberger *et al.* 1988) and *Eco*R I (McClarin *et al.* 1986). [An interesting variation involving

water-mediated hydrogen bonding in the major groove is seen in the *trp* repressor complex with DNA (Otwinowski *et al.* 1988).] In general, structural complementarity between a protein and a specific DNA sequence is achieved in idiosyncratic manners: there does not appear to be a code for nucleic-acid sequence recognition (Pabo, 1983; Matthews, 1988). While particular amino-acid side chains do not always recognize the same base pair, there are some apparent preferences as suggested by Seeman *et al.* (1976). The guanidinium group of arginine often makes a bidentate interaction with the N7 and O6 of guanine (observed with *Eco*R I (McClarin *et al.* 1986) and *Trp* repressor (Otwinowski *et al.*, 1988) and proposed for *λcro* repressor (Ohlendorf, *et al.* 1982) and CAP (Weber and Steitz, 1984). However, it is also observed to interact with the N7s of two adjacent adenines in *Eco*R I (McClarin *et al.*, 1986). Similarly, while the hydrogen-bond donors and acceptors of the glutamine side chain are observed frequently to interact with the corresponding hydrogen-bond donors and acceptors on adenine (Aggarwal *et al.* 1988; Jordan & Pabo, 1988), other interactions of glutamine are also seen, such as the interaction of its NH_2 with the O6 and N7 of guanine in 434 repressor. Perhaps somewhat unexpectedly the carboxylate side chain of glutamic acid is observed in the case of *Eco*R I (Frederick *et al.* 1985) and proposed in the case of CAP (Weber & Steitz, 1984) to simultaneously interact with two adjacent base pairs (the N6s of two adenines in *Eco*R I). The ability of these side chains to make bidentate interactions with DNA greatly enhances their suitability for sequence-specific recognition (Seeman *et al.* 1976). The van der Waals interactions between the protein and the 5-methyl group of thymine appear also to contribute to specificity. Presumably, the close packing of protein against a GC base pair would, in many cases, sterically exclude its replacement by an AT base pair with its accompanying bulky 5-methyl group.

It has frequently been proposed that the sequence specificity of a protein might be altered by changing a side chain such as arginine that recognizes the guanine in a GC base pair for a side chain such as glutamine that might recognize the adenosine in an AT base pair. Close examination of the detailed crystal structures of DNA complexes as well as simple model building suggests that this is not likely to work. First of all, the glutamine and arginine side chains are of different lengths and are thus not strictly interchangeable, other features of protein and nucleic acid being held constant. Secondly, those protein side chains interacting with the bases can be involved in a more extensive network of interactions within the protein allowing the formation of a complex protein surface. Thus, for example, Glutamine 44 in λ repressor (Jordon & Pabo, 1988) is interacting both with an adenosine and with glutamine 33 that in turn is also interacting with the backbone phosphate.

2.3 *Role of nucleic acid bendability*

The extent to which the sequence-dependent bendability or deformability of DNA would play a vital role in nucleic-acid sequence recognition was perhaps less

well anticipated. It appears, however, to play an extremely important role in most though perhaps not all of the complexes. Richmond & Steitz (1976) concluded that '...sequence-dependent alteration of the double-stranded DNA conformation are probably important in *lac* repressor specificity' in order to account for the two to three orders of magnitude difference in *lac* repressor affinity for poly[d(AUHgX)] with various different bulky substituents at the 5 position of U. Crick & Klug (1975) hypothesized that the wrapping of DNA around the histone core in the nucleosome might be achieved by periodic sharp bends or 'kinks' rather than smooth bending. In spite of a few early hints no direct evidence for the role of DNA distortion was obtained until co-crystal structures of complexes were determined.

Significant distortion of the DNA or RNA structure from its presumed structure in solution is observed in the crystal structures of DNA complexes with *Eco*R I (Frederick *et al.* 1985; McClarin *et al.* 1986), 434 repressor (Aggarwal *et al.* 1988), *trp* repressor (Otwinowski *et al.* 1988), DNAse–I (Suck *et al.* 1988), Klenow fragment (Steitz *et al.* 1987; Freemont *et al.* 1988) and the tRNA complex with GlnRS (Rould *et al.* 1989). The structure of the nucleosome although at low resolution shows significant kinking in the DNA (Richmond *et al.* 1984). Model building (Warwicker *et al.* 1987) and the recent determined co-crystal structure (Schultz, Shields and Steitz, unpublished) as well as DNA binding and gel shift experiments (Wu & Crothers, 1984; Gartenberg & Crothers, 1988) establish an important role for DNA kinking and bendability in the sequence-specific binding of CAP to DNA, as delineated in detail below.

The distortions of duplex DNA structure that have been observed in complexes include changes in twist (Aggarwal *et al.* 1988), groove width (Suck *et al.* 1988) and kinks (McClarin *et al.* 1986; Otwinowski *et al.* 1988; Aggarwal *et al.* 1988; Warwicker *et al.* 1987; Schultz, Shields & Steitz 1990, unpublished). The two types of kinks observed thus far are an abrupt reduction in the twist between adjacent base pairs which widens the major groove (McClarin *et al.* 1986) and a change in the roll-angle between successive base pairs resulting in a bend in the DNA helix axis. As initially described by Drew & Travers (1984) AT-rich sequences favour bending into the minor groove while GC-rich sequences facilitate kinks that narrow the major groove.

The sequence-dependent deformability of duplex DNA or RNA that provides specificity for sequences being recognized by the protein can include the melting of base pairs. If binding to protein requires melting of one or more base pairs, then the binding of mismatched base pairs should be favoured over AT pairs that in turn should bind better than GC pairs. The order of binding should reflect the thermodynamic stability of the base pairs.

Two examples of the role of duplex meltability in sequence specificity can be cited. Binding of $tRNA^{Gln}$ to its cognate synthetase results in the breaking of the base pair between nucleotides U_1 and A_{72} (Rould *et al.* 1989). For GlnRS recognition and charging of tRNA it is important that this base pair be not GC (Yarus *et al.* 1977). Presumably, the added free energy cost of breaking a GC pair

makes tRNAs containing a GC at 1–72 less suitable for proper binding to the enzyme.

The 3′5′-exonuclease-active site of *E. coli* DNA polymerase I is observed to denature duplex DNA and bind four single-stranded nucleotides at the 3′ terminus (Steitz *et al.* 1987; Freemont *et al.* 1988). The physiological role of this exonuclease is to 'edit' out mismatched base pairs that are erroneously incorporated at the polymerase active site some 30 Å away. It has been proposed that the specificity of this exonuclease for editing out mismatch base pairs with higher frequency than correctly matched base pairs arises from the former's lower thermal stability (Brutlag & Kornberg, 1972; Steitz *et al.* 1987). In the competition between the duplex-binding polymerase active site and the single-strand binding exonuclease-active site, duplex DNA containing mismatched base pairs will bind to the exonuclease site with greater frequency than correctly matched duplex.

Sequence recognition also arises from the sequence-dependent ability of single-stranded nucleic acid to take up the conformation required for protein binding. The single-stranded acceptor end of $tRNA^{Gln}$ assumes a hairpinned conformation upon binding to GlnRS (Rould *et al.* 1989). This conformation is stabilized by an interaction between the N2 of G_{73} and the backbone phosphate of A_{72}. the inability of the other three bases to form this stabilizing hydrogen bond would be expected to increase the free energy cost of binding $tRNA^{Gln}$ with other bases at this position.

Nucleic-acid distortability as a more indirect source of sequence specificity arises from two facts: (1) Proteins often bind a conformation of a nucleic acid that is altered from its uncomplexed solution conformation. (2) The free energy cost for various nucleic-acid sequences to assume the conformation that is required for its binding to the protein is not the same for different sequences.

2.4 *Role of water molecules in sequence recognition*

Two protein–DNA complexes and one protein–RNA complex are sufficiently well refined at high resolution to show the positions of firmly bound water molecules. Some buried water molecules appear to be playing an important role in both DNA and RNA sequence recognition. Water (or a serine hydroxyl group) can only make a base-specific hydrogen bond if it is also making at least two other hydrogen bonds with obligate donors or acceptors on the protein and is buried form bulk solvent. In this circumstance the two unsatisfied water H-bond donor/acceptors directed towards the nucleic acid become obligate donor/acceptors and become part of the H-bonding template surface of the protein to which the nucleic acid must be complementary for optimal binding. In the trpR–DNA complex there are three water molecules per half operator bound between protein and the DNA bases and two of them appear to be making hydrogen bonds that specify base pairs 5, 6 and 7 from the dyad axis (Otwinowski *et al.* 1988). In this case, water molecules are playing the role of 'honorary' protein side chains. In the 434

repressor complex a water molecule is used or orient the carboxy amide side chain of Gln33 so that the amide is direct towards the O4 of thymine (Aggarwal *et al.* 1988). Water molecules also mediate the interaction of the guanidinium group of Arg43 with AT base pairs in the minor groove.

In the GlnRS complex with tRNA a buried water molecule is an integral part of the H-bonding matrix presented to the shallow groove of the tRNA acceptor stem (Rould *et al.* 1989). Hydrogen bonds between this water molecule and both a buried carboxylate of Asp235 and the backbone amide of residue 183 serve to orient one hydrogen bond donor of the water towards the O2 of cytosine 71 and an acceptor towards the N2 of guanine 2.

2.5 *Role of the minor groove in DNA and RNA recognition*

As pointed out by Seeman *et al.* (1976), thee are fewer features presented by base pairs in the minor groove (as compared with the major groove) that allow discrimination among the two base pairs and their two orientations (Fig. 2). The hydrogen-bond acceptors (N3 on guanine and adenine and O2 on cysteine and thymine) occur in almost identically the same place in the minor groove for all four bases. Only the exocyclic N2 of guanine distinguishes AT from GC and perhaps GC from CG. Furthermore, the minor groove of B-DNA is too narrow to accommodate an α-helix and too deep for the bases to be reached by side chains alone.

It appears, however, that there are ways in which interactions in the minor groove can be sequence specific. The water mediated interactions made by the two symmetry related Arg43 guanidinium groups in 434 repressor stabilize a highly propeller-twisted structure taken up by an AT-rich sequence at the dyad axis of the 434 operator (Aggarwal *et al.* 1988). Since a GC-rich sequence cannot take up this structure it presumably would not be able to make these water-mediated contacts with the protein and would bind less tightly than the AT-rich sequence. In this case sequence specificity is achieved by both water-mediated hydrogen-bonding complementarity and sequence-dependent deformability of DNA.

The sequence preferences exhibited in the DNAse I cleavage of DNA arise from its interactions in the minor groove (Suck *et al.* 1988). The side chain of Tyr that is observed to bind in the minor groove will fit into the normal width minor groove, but not into the narrower minor groove that characterizes AT-rich sequences.

With duplex RNA, which is A-form, the accessibility of the grooves is the reverse of DNA; only the minor or shallow groove is accessible, the major groove being both narrow and very deep. Two sequence-specific interactions between GlnRS and the minor groove of $tRNA^{Gln}$ have been observed (Rould *et al.* 1989). The carboxylate of an aspartic-acid side chain emanating from the amino end of an α-helix interacts simultaneously with the N2 of guanine, the 2′OH of a backbone ribose, and a buried water molecule. One would anticipate that substitution of guanine by any of the other three bases would abolish this

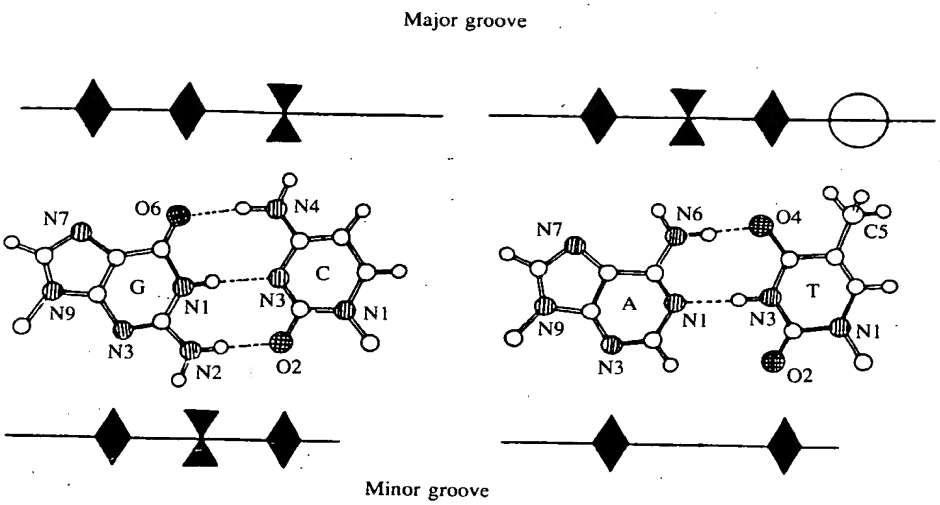


Fig. 2. The hydrogen-bond donors and acceptors presented by Watson–Crick base pairs to the major groove and the minor groove and the minor groove (adapted from Lewis *et al.* 1985). The symbols for hydrogen bond donors (⧗) and acceptors (◆) (Woodbury *et al.* 1989) show a varied pattern presented by the base pairs to the major and an poor information array in the minor groove. While it is possible to distinguish among AT, TA, GC and CG in the major groove, functional groups in the minor groove allow only easy discrimination between AT and GC containing base pairs. ○, methyl group.

interaction thus lowering the affinity, a proposal that is currently being tested. This particular sequence-specific bidentate interaction can only occur with RNA. Another sequence dependent interaction occurs between the peptide backbone carbonyl oxygen of proline and the N2 of quanine, again in the minor groove.

3. DNA-BINDING STRUCTURE MOTIFS

The structures of three different motifs that interact with DNA in a sequence-specific manner are now known: the helix–turn–helix motif characteristic of most prokaryotic regulatory proteins, α-helical structure found in *Eco*R I, and the zinc 'finger' structure found in some eukaryotic transcription factors. In order for a protein to present a surface that is complementary to the major groove of B-DNA, it is necessary for the protein structure to protrude significantly from its own surface. In the case of all three of these motifs it appears likely that an α-helix is utilized in order to present the array of side chains that are complementary to the exposed edges of base-pairs in the major groove of B-DNA. It has been recognized for 30 years that an α-helix will fit into the major groove of B-DNA (Zubay & Doty, 1959) and early models for histone interaction with DNA speculated that an α-helix might fit into the major groove (Sung & Dixon, 1970). In these and later models for *lac* repressor interaction with DNA (Adler *et al.* 1972) and protamine

AA

Form Approved Through 9/30/94
OMB No. 0925-0001

DEPARTMENT OF HEALTH AND HUMAN SERVICES
PUBLIC HEALTH SERVICE

GRANT APPLICATION

Instructions carefully. Typ in the [redacted] areas only.
Type density must be 10

LEAVE BLANK FOR PHS USE ONLY.		
Type	Activity	Number
Review Group		Formerly
Council/Board *(Month, Year)*		Date Received

1. TITLE OF PROJECT (Do not exceed 56 typewriter spaces.)
Zinc Finger Proteins as anti-HIV Therapeutics

2a. RESPONSE TO SPECIFIC REQUEST FOR APPLICATIONS OR PROGRAM ANNOUNCEMENT ☒ NO ☐ YES *(If "YES," state number and title)*
Number: Title:

OTPE 4C95 MAY 0 5 2003 PATENT & TRADEMARK OFFICE

2b. TYPE OF GRANT PROGRAM R01

3. PRINCIPAL INVESTIGATOR/PROGRAM DIRECTOR

3a. NAME *(Last, first, middle)*
Barbas III, Carlos F.

3b. DEGREE(S)
Ph.D.

3c. SOCIAL SECURITY NO.
261-77-3269

3d. POSITION TITLE
Assistant Member

3e. MAILING ADDRESS *(Street, city, state, zip code)*
10666 No. Torrey Pines Road
La Jolla, CA 92037

3f. DEPARTMENT, SERVICE, LABORATORY, OR EQUIVALENT
Molecular Biology

3g. MAJOR SUBDIVISION

3h. TELEPHONE AND FAX *(Area code, number and extension)*
TEL: (619) 554-9098
FAX: (619) 554-6778

BITNET/INTERNET ADDRESS

4. HUMAN SUBJECTS
☒ 4a. NO ☐ YES If "Yes," exemption no. or IRB approval date 4b. Assurance of compliance no.

5. VERTEBRATE ANIMALS
☒ 5a. NO ☐ YES If "Yes," IACUC approval date 5b. Animal welfare assurance no.

6. DATES OF ENTIRE PROPOSED PROJECT PERIOD		7. COSTS REQUESTED FOR INITIAL BUDGET PERIOD		8. COSTS REQUESTED FOR ENTIRE PROPOSED PROJECT PERIOD	
From (MMDDYY)	Through (MMDDYY)	7a. Direct Costs ($)	7b. Total Costs ($)	8a. Direct Costs ($)	8b. Total Costs ($)
120195	113098	320,203	498,302	875,157	1,434,081

9. PERFORMANCE SITES *(Organizations and addresses)*
The Scripps Research Institute
Dept. of Molecular Biology
10666 No. Torrey Pines Road
La Jolla, CA 92037

10. INVENTIONS AND PATENTS *(Competing continuation application only)*
☐ NO ☐ YES If "YES," ☐ Previously reported ☐ Not previously reported

11. NAME OF APPLICANT ORGANIZATION
The Scripps Research Institute
ADDRESS 10666 No. Torrey Pines Road
La Jolla, CA 92037

12. TYPE OF ORGANIZATION
☐ Public: *Specify* ☐ *Federal* ☐ *State* ☐ *Local*
☒ Private Nonprofit
☐ Forprofit *(General)* ☐ Forprofit *(Small Business)*

13. ENTITY IDENTIFICATION NUMBER
33-0435954

Congressional District
41

14. BIOMEDICAL RESEARCH SUPPORT GRANT CREDIT
Code: 60 Identification: Other Research Org.

15. NAME OF ADMINISTRATIVE OFFICIAL TO BE NOTIFIED IF AWARD IS MADE
Kaye I. Wynne
TELEPHONE (619) 554-8653
FAX (619) 554-6152
TITLE Director, Grants and Contracts
ADDRESS 10666 No. Torrey Pines Road
La Jolla, CA 92037

BITNET/INTERNET ADDRESS

16. NAME OF OFFICIAL SIGNING FOR APPLICANT ORGANIZATION
William H. Beers, Ph.D.
TELEPHONE (619) 554-9792
FAX (619) 554-9899
TITLE Sr. Vice President, TSRI
ADDRESS 10666 No. Torrey Pines Road
La Jolla, CA 92037

BITNET/INTERNET ADDRESS

17. PRINCIPAL INVESTIGATOR/PROGRAM DIRECTOR ASSURANCE: I agree to accept responsibility for the scientific conduct of the project and to provide the required progress reports if a grant is awarded as a result of this application. Willful provision of false information is a criminal offense (U.S. Code, Title 18, Section 1001). I am aware that any false, fictitious, or fraudulent statement may, in addition to other remedies available to the Government, subject me to civil penalties under the Program Fraud Civil Remedies Act of 1986 (45 CFR 79).

SIGNATURE OF PERSON NAMED IN 3a. *(In ink. "Per" signature not acceptable.)*

DATE 4/14/95

18. CERTIFICATION AND ACCEPTANCE: I certify that the statements herein are true and complete to the best of my knowledge, and accept the obligation to comply with Public Health Service terms and conditions if a grant is awarded as the result of this application. A willfully false certification is a criminal offense (U.S. Code, Title 18, Section 1001). I am aware that any false, fictitious, or fraudulent statement may, in addition to other remedies available to the Government, subject me to civil penalties under the Program Fraud Civil Remedies Act of 1986 (45 CFR 79).

SIGNATURE OF PERSON NAMED IN 16. *(In ink. "Per" signature not acceptable.)*

DATE 4/18/95

DESCRIPTION: State the application's broad, long-term objectives and specific aims, making reference to the health relatedness of the project. Describe concisely the research design and methods for achieving these goals. Avoid summaries of past accomplishments and the use of the first person. This abstract is meant to serve as a succinct and accurate description of the proposed work when separated from the application. **DO NOT EXCEED THE SPACE PROVIDED.**

This research aims to develop a new class of therapeutic proteins. Zinc finger proteins will be evolved to work as defined genetic switches to turn-on or turn-off gene expression on demand. The primary goal is to develop these proteins as specific inhibitors of HIV-1 replication which act at the level of transcription. A cocktail of high affinity HIV-1 specific DNA binding proteins will be produced. These proteins will recognize highly conserved sequences in the genome of HIV-1, bind to them, and stop viral transcription. Transcription will be inhibited by the direct administration of soluble protein. Reduced levels of viral replication may have a profound effect in combination with already existing drugs which inhibit *de novo* infection. Alternatively, the proteins could be utilized in a gene therapy approach to produce cell populations which do not support viral replication.

In vitro approaches will be developed to modify the specificity and affinity of an existing zinc finger protein, Zif268. Preliminary studies demonstrate that using the phage display approach, both the sequence specificity and affinity of this protein for DNA may be changed. Modularity of the zinc finger domain has been demonstrated by tethering together as many as six selected domains to create new proteins which recognize up to 18 base pairs of DNA sequence. This approach will be further refined to create 4 proteins which bind conserved sites in the HIV-1 genome. Proteins will be characterized with respect to affinity, sequence specificity, and ability to inhibit transcription of reporter genes and replication of laboratory and primary clinical isolates of HIV-1. The utility of known cellular uptake and nuclear targeting sequences to deliver extracellular zinc finger protein from the medium to the nucleus will be examined. Selection and screening experiments will be performed to optimize the efficiency of these sequences as well as to attempt to discover new ones. This strategy will allow the protein itself to be used as a drug to selectively modulate transcription and would avoid the use of a gene therapy approach to delivery. To investigate their application in a gene therapy approach, stable cell lines which express the evolved proteins will be constructed and their ability to support viral replication examined. The result of this work will be a cocktail of evolved zinc finger proteins which are effective when delivered as soluble proteins or when expressed within a cell and act to block HIV-1 transcription and replication by sequence-specific recognition of conserved regions in the HIV-1 genome. This research will provide proteins for therapeutic application in HIV-1 infection and generate of a collection of zinc finger domains which can be rapidly assembled to recognize any given stretch of DNA sequence. NMR studies of these proteins in complex with DNA will provide insight into the molecular basis for sequence-specific recognition of DNA. The success of this project may have implications for the treatment of many human ailments.

PERSONNEL ENGAGED ON PROJECT, INCLUDING CONSULTANTS/COLLABORATORS. *Use continuation pages as needed* to provide the required information in the format shown below on ***all*** individuals participating in the scientific execution of the project.

Name Carlos F. Barbas, III	Degree(s) Ph.D.	Social Security No. 261-77-3269
Position Title Assistant Member	Date of Birth (MM/DD/YY) 11/05/64	Role on Project P.I.
Organization The Scripps Research Institute		Department Molecular Biology
Name Herren Wu	Degree(s) Ph.D.	Social Security No. 024-70-2425
Position Title Research Associate	Date of Birth (MM/DD/YY) 12/14/63	Role on Project Postdoc.
Organization The Scripps Research Institute		Department Molecular Biology
Name Qiang Liu	Degree(s) Ph.D.	Social Security No. 479-173-3787
Position Title Research Associate	Date of Birth (MM/DD/YY)	Role on Project Postdoc.
Organization The Scripps Research Institute		Department Molecular Biology
Name Serge Bergeron	Degree(s) B.S.	Social Security No. 606-28-6999
Position Title Graduate Student	Date of Birth (MM/DD/YY) 10/2/68	Role on Project Graduate Student
Organization The Scripps Research Institute		Department Molecular Biology
Name Sally Pinz-Sweeney	Degree(s) B.S.	Social Security No. 555-29-8949
Position Title Research Technician II	Date of Birth (MM/DD/YY) 8/11/64	Role on Project Research Tech. II
Organization The Scripps Research Institute		Department Molecular Biology
Name Flossie Wong-Staal	Degree(s) Ph.D.	Social Security No.
Position Title Professor	Date of Birth (MM/DD/YY)	Role on Project Collaborator
Organization University of California, San Diego		Department Medicine/Biology
Name Joel Gottesfeld	Degree(s) Ph.D.	Social Security No.
Position Title Associate Member	Date of Birth (MM/DD/YY)	Role on Project Collaborator
Organization The Scripps Research Institute		Department Molecular Biology

Additional Personnel Engaged on Project:

Name: John P. Moore Degree(s): Ph.D. Social Security No. 055-80-0806

Position Title: Staff Investigator Date of Birth(MM/DD/YY): 01/20/57 Role on Project: Collaborator

Organization: New York University School of Medicine Department: Microbiology

Name: Joseph Sodroski Degree(s): Ph.D. Social Security No. .

Position Title: Associate Professor Date of Birth(MM/DD/YY): Role on Project: Collaborator

Organization: Harvard Medical School Department: Pathology

Name: Peter Wright Degree(s): Ph.D. Social Security No.

Position Title: Dept. Chairman Date of Birth(MM/DD/YY): Role on Project: Collaborator

Organization: The Scripps Research Institute Department: Molecular Biol.

Name: Frederic D. Bushman Degree(s): Ph.D. Social Security No.

Position Title: Assistant Professor Date of Birth(MM/DD/YY): Role on Project: Collaborator

Organization: The Salk Institute for Biological Studies Department: Biological Stud.

Name: Degree(s): Social Security No.

Position Title: Date of Birth(MM/DD/YY): Role on Project:

Organization: Department:

Name: Degree(s): Social Security No.

Position Title: Date of Birth(MM/DD/YY): Role on Project:

Organization: Department:

Type the name of the principal investigator/program director at the top of each printed page and each continuation page. (For type specifications, see **Specific Instructions** on page 10.)

RESEARCH GRANT
TABLE OF CONTENTS

*Type density and size must conform to limits provided in Specific Instructions on page 10.

Appendix *(Five collated sets. No page numbering necessary for Appendix)*

Number of publications and manuscripts accepted or submitted for publication *(Not to exceed 10)* 3

Other items (list):

[X] Check If Appendix is Included

DETAILED BUDGET FOR INITIAL BUDGET PERIOD
DIRECT COSTS ONLY

FROM 12/01/95 THROUGH 11/31/96

PERSONNEL (Applicant organization only) NAME	ROLE ON PROJECT	TYPE APPT. (months)	% EFFORT ON PROJ.	INST. BASE SALARY	DOLLAR AMOUNT REQUESTED (omit cents) SALARY REQUESTED	FRINGE BENEFITS	TOTALS
Carlos F. Barbas, III	Principal Investigator	12	60		50,639	10,634	61,273
Herren Wu	Research Assoc.	12	100		27,310	5,189	32,499
Qiang Liu	Research Assoc.	12	100		26,000	4,940	30,940
Serge Bergeron	Graduate Student	12	100		16,000	3,040	19,040
Sally Pinz-Sweeney	Research Tech. II	12	100		35,516	6,748	42,264
SUBTOTALS →					155,465	30,551	186,016

CONSULTANT COSTS

N/A — Ø

EQUIPMENT (Itemize) A 7% sales tax is included with the list price.
Storage phosphorimager screen & cassette ($1,177)
Wallac Berthod automatic luminometer ($30,174)
for 96 well microplates with temperature control
Millipore DNA/RNA synthesizer with trityl monitor ($29,425) — 60,776

SUPPLIES (Itemize by category)
Pipettor ($637)
Opague luminometer plates ($438)
Luciferase/SEAP assay rgts. ($1400)
Tissue culture ($5000)
Microbiological supplies ($7200)
ELISA reagents/supplies ($925)
p24 ELISA kits ($3600)
Streptavidin, 10mg ($430)
BIAcore sensor chips/reagents ($5200)
Purification materials ($3600)
Radioisotopes ($1200)
Restriction enzymes ($6000)
PCR enzymes/tubes/reagents ($4200)
DTT, 100g ($581), Agarose ($750)
Sequencing ($10000), film ($300)
Nucleic acid synthesis supp. ($6000)
MW markers, plasmids ($1000)
Plastic, glassware ($4800)
— 63,261

TRAVEL
One meeting for P.I. per year — 1,200

PATIENT CARE COSTS
INPATIENT — Ø
OUTPATIENT — Ø

ALTERATIONS AND RENOVATIONS (Itemize by category)
N/A — Ø

OTHER EXPENSES (Itemize by category) Tuition ($5000), glasswashing ($1200), publication costs & reprints ($1000), photo/artwork ($1,000), journals ($500), shipping ($250) — 8,950

SUBTOTAL DIRECT COSTS FOR INITIAL BUDGET PERIOD — 320,203

CONSORTIUM/CONTRACTUAL COSTS
DIRECT COSTS $
INDIRECT COSTS $
TOTAL → Ø

TOTAL DIRECT COSTS FOR INITIAL BUDGET PERIOD (Item 7a, Face Page) → $ 320,203

BUDGET FOR ENTIRE PROPOSED PROJECT PERIOD
DIRECT COSTS ONLY

BUDGET CATEGORY TOTALS		INITIAL BUDGET PERIOD *(from page 4)*	ADDITIONAL YEARS OF SUPPORT REQUESTED 2nd	3rd	4th	5th
PERSONNEL: *Salary and fringe benefits Applicant organization only*		186,016	193,458	201,196		
CONSULTANT COSTS		---	---	---		
EQUIPMENT		60,776	---	---		
SUPPLIES		63,261	67,057	71,080		
TRAVEL		1,200	1,272	1,348		
PATIENT CARE COSTS	INPATIENT	---	---	---		
	OUTPATIENT	---	---	---		
ALTERATIONS AND RENOVATIONS		---	---	---		
OTHER EXPENSES		8,950	9,487	10,056		
SUBTOTAL DIRECT COSTS		320,203	271,274	283,680		
CONSORTIUM/ CONTRACTUAL COSTS		---	---	---		
TOTAL DIRECT COSTS		320,203	271,274	283,680		

TOTAL DIRECT COSTS FOR ENTIRE PROPOSED PROJECT PERIOD *(Item 8a)* ⟶ $ 875,157

JUSTIFICATION (Use continuation pages if necessary):

From Budget for Initial Period: Describe the specific functions of the personnel, collaborators, and consultants and identify individuals with appointments that are less than full time for a specific period of the year, including VA appointments.

For All Years: Explain and justify purchase of major equipment, unusual supplies requests, patient care costs, alterations and renovations, tuition remission, and donor/volunteer costs.

From Budget for Entire Period: Identify with an asterisk (*) on this page and justify any significant increase or decrease in any category over the initial budget period. Describe any change in effort of personnel.

For Competing Continuation Applications: Justify any significant increases or decreases in any category over the current level of support.

(see following pages)

BARBAS III, C.F.

JUSTIFICATION OF PERSONNEL, EQUIPMENT, AND SUPPLIES

Personnel

Carlos F. Barbas III, Ph.D. 60%. Dr. Barbas is an Assistant Member in the Department of Molecular Biology at the Scripps Research Institute and will serve as the principal investigator. He has pioneered the use of phage display for the construction and selection of combinatorial antibodies from immune sources and the design of synthetic antibody libraries for the production of antibodies without immunization. I have devoted considerable efforts to create systems and strategies for the manipulation of antibody specificity and affinity in hopes of applying human antibodies in the service of human health. Extensive work in the area of protein engineering has allowed us to modify zinc finger proteins with the same ease we manipulate antibodies. The ability to produce DNA binding proteins at will to any given target sequence should have tremendous applications to human health as well as basic science. This area is becoming and will be our major focus in the years to come. The request for 60% salary reflects this emphasis. Scripps requires investigators to cover 100% of their salary on grants and 60% reflects my effort and commitment to the project. I will be responsible for experimental design, implementation of the proposed work, and coordination of experiments with collaborators. I will be active in the experimental work with a focus on aims 2,3, and 4.

Herren Wu, Ph.D. 100%. Dr. Wu is a molecular biologist who earned his Ph.D. in Biochemistry from University of Massachusetts in 1993 were he studied protein recognition of RNA. He has been working in the P.I.'s group since the fall of 1993. Since this time he has become an expert in library construction and the phage display approach. Furthermore, he has gained extensive experience with the Pharmacia BIAcore instrument for the study of zinc finger/DNA interactions. He has been instrumental in getting the zinc finger project going and has extensive experience in characterizing protein/nucleic acid interactions. His major focus will be aim 1.

Qiang Liu, Ph.D. 100%. Dr. Liu is a molecular biologist who earned his Ph.D. in 1995 from Duke University. He has been working in the P.I.'s group since March of 1995. He has extensive experience in the study of eukaryotic promoters and their regulation in the retinoic acid system. His focus will be the development of reporter gene assays in tissue culture and the development and optimization of uptake sequences which will allow the protein to be used as a drug directly. He will characterize the ability of evolved proteins to enhance and repress transcription and will be involved in assays of the proteins ability to inhibit HIV. Thus his focus will be aims 2 and 3.

Serge Bergeron, Graduate Student 100%. Mr. Bergeron has been in my group for the last 6 months learning molecular biology techniques associated with this project. His primary focus will be in the manipulation of finger specificity, affinity, and connectivity.

Barbas (budget justification, pg. 2)

He will be trained in the area of protein engineering and the study of protein/nucleic acid interactions. He will work closely with postdoctoral fellows. He will develop a dissertation which focuses on understanding zinc finger protein interactions with DNA and will be focused on aim 1. Mr. Bergeron received his B.S. degree from McGill University in 1994.

Sally Pinz-Sweeney, B.S 100%. She has been a technician in my group for the three years. She is skilled in protein purification, DNA sequencing, library construction, tissue culture, and basic molecular biology. She prepares many of the reagents which we use saving us thousands of dollars over commercially available products, for example large scale preparation of ultracompetent cells for electroporation. She will closely with the P.I. and will assist others in expression, ELISA characterization, purification, sequencing, and the preparation of microbiological supplies.

Collaborators

A team of exceptional collaborators has been assembled to examine every possible application of evolved zinc fingers as anti-HIV therapeutics and to enhance our understanding of these proteins at the molecular level. This team has the "critical mass" that will lay the foundation which will lead us to human testing in the most expeditious way.

Frederic Bushman, Ph.D. is an assistant Professor at the Salk Institute located down the road from Scripps. His research is directed toward understanding the mechanism of HIV integration. He is a recognized expert on HIV integration. His laboratory will perform and assist in performing assays to explore the ability of evolved zinc fingers to inhibit HIV integration. We will provide him with evolved zinc finger proteins to various sites in the HIV genome to assist his studies in defining the integration complex. As his understanding of integration evolves prospects for the development of new zinc finger targets will also evolve and we will work together with his laboratory to explore new targets as they develop.

John Moore, Ph.D. is an Associate Professor at the Aaron Diamond AIDS Research Center. He has an international reputation in the AIDS research community. He will perform neutralization assays with primary clinical isolates of HIV-1 and examine reduction in viral output when infected cells are incubated with evolved zinc fingers. We will also provide him with stable cell lines expressing evolved zinc fingers and he will examine their resistance to infection with a number of HIV isolates. He will also attempt to select for escape variants with cocktails of evolved zinc fingers. We currently maintain a very productive collaboration on human anti-HIV antibodies.

Barbas (budget justification, pg. 3)

Flossie Wong-Staal, Ph.D. of UCSD has been an important AIDS researcher from the start. A large portion of her recent efforts are targeted to the development of ribozymes as an intracellular immunization strategy. Our zinc finger project should be complementary to her ongoing ribozyme work. Her laboratory is across the road from Scripps so she serves as an important proximal resource. If the zinc finger approach develops along the lines of a gene therapy strategy she will assist us along this path. She will assess the ability of zinc finger proteins to work together with her ribozymes to inhibit HIV.

Joel Gottesfeld, Ph.D. is a member of the Scripps Research Institute Department of Molecular Biology. Dr. Gottesfeld is an expert in gene expression and the study of protein/nucleic acid interactions. His laboratory was the first to characterize zinc finger RNA interactions and determine the fingers in TFIIIA responsible for binding to RNA. Recently, he has mapped the binding of TFIIIA to 5SRNA. Dr. Gottesfeld serves as an important resource with many years experience in the area of transcription factors and shares a common interest with the P.I. in understanding and exploiting zinc finger proteins for human health purposes. He has studied the ability of the three fingers of TFIIIA to inhibit transcription, see attached letter. He has also exploited the repressor activity of the three DNA binding zinc fingers of TFIIIA to inhibit RNA pol III transcription and dissect the transcription complex.

Joseph Sodroski, Ph.D. is an associate professor of the Harvard Medical School and is located at the Dana-Farber Cancer Institute. He will perform quantitative assays to examine the effect of our proteins on HIV-1 replication. He will also examine their effects on a number of reporter gene constructs in his laboratory. We currently maintain a very productive collaboration on human anti-HIV antibodies.

Peter Wright, Ph.D. is a member and Chairman of the Scripps Research Institute Department of Molecular Biology. In recent years a major focus of his research has been NMR structural determination of zinc finger proteins in complex with DNA. Dr. Wright's laboratory was the first to publish a high resolution solution structure of a zinc finger in 1989, and is now refining the solution structure of the first three zinc fingers of TFIIIA in complex with DNA. Attempts by many labs to crystallize this complex have all failed. His laboratory has one of the few high resolution 750MHz NMR instruments in the world and he is recognized world-wide as a leader in the use of NMR in the determination of protein structure. He is highly committed to the determining how these fingers function and is currently funded by NIH GM36643 to elucidate structures of zinc finger proteins.

Equipment: We request funds to obtain a DNA/RNA synthesizer. This project is heavily dependent on oligonucleotides to construct and select zinc finger libraries and to thoroughly characterize their binding characteristics. Library construction requires specifically doped oligonucleotides. DNA is prepared specifically labeled with biotin or

an amino group to allow for its immobilization to proteins, supports, and ELISA. Labeled DNA is used extensively in the ELISA protocols we have developed to characterize binding and to standardize purified proteins. This technique allows us not to be heavily dependent on radioisotopes but means that we utilize a lot of oligo, thus a number of oligonucleotides will need to be synthesized several times. These special requirements driveup the cost of these reagents if they are purchased from commercial sources. We have requested $6000 for nucleic acid synthesis reagents and solvents. This amount would be almost tripled if oligos are purchased commercially as we have been doing. Thus the instrument will more than pay for itself in the first three years.

The second instrument we request funds for is a luminometer . This instrument will be used extensively in reporter gene assays with both luciferace and secreted alkaline phosphatase constructs. We will test the inhibitory as well as enhancing function of our proteins using this instrument. The instrument will be extensively utilized to screen large numbers of proteins for efficient uptake from the media and nuclear localization. This will allow us to define optimal uptake sequences and will utilize the high-throughput capable with the microtiter format we request. The last piece of equipment is a phosphorimager plate. We have access to a departmental instrument. This will allow us to save a considerable amount of time and money for film. For binding studies and transcriptions assays this is very useful but for DNA footprinting we utilize film.

Materials and Supplies: This project involves extensive use of oligonucleotides and PCR for the construction of numerous libraries. Reporter constructs need to be modified using PCR or site directed mutagenesis. As a result the project requires extensive use of restriction enzymes and DNA modification enzymes, the basic tools for recombinant DNA work. To verify library quality and to quickly characterize mutant proteins following selection using phage display, sequencing of mutants is required. A large portion of the supply costs are incurred because the project will create and examine the function of many different DNA binding proteins. Characterization of numerous zinc-finger variants requires DNA sequencing, ELISA, chromatography, gel retardation assays, and BIAcore analysis. DNA synthesis and sequencing are major cost. For synthesis we request $6000 which reflects the large number of oligos which must be produced for library construction, selection, specificity characterization, and sequencing as well as special labeling with amino groups and biotin. Utilizing the core sequencing facility in our department we save a considerable amount of personnel time and money. The preparation and each sequence however is $25 and this project is very heavy on sequencing of constructs and individual clones. We request $10,000 for sequencing expenses allowing us to sequence 200 constructs in both directions. Note the project will generate 64 optimal proteins in a single specific aim and the generation of each one of the 64 will require sequencing of as many as 10. Our demonstrated ability to produce and characterize antibody variants supports our ability to do the same in this system. Further, aach library is characterized by sequencing prior to selection. Rapid ELISA based screens allow us to characterize binding characteristics in a semi-quantitative fashion. For this we a anti-HA -antibody-alkphos conjugate we

prepare in the lab. The libraries we construct in phage display vectors require a considerable amount of PCR and restriction and ligation enzymes. For PCR we request $4200 or $350 per month which includes a TA cloning kit every three months. For restriction and other enzymes we request $6000 or $500 per month. For plastics and glassware which covers filtration devices, centracon concentrators, pipette tips, and tubes as well except for PCR we request $4800. To avoid contamination in phage display we utilize barrier tips (10 boxes cost $91) and many disposable tubes. For reporter gene assays and optimization of uptake sequences we require $5000 for tissue culture which covers culture ware and media, and fetal calf sera. We will also produce our own anti-HA antibody, which we utilize to detect HA-tagged proteins; in tissue culture . This saves a considerable amount of money since the anti-HA antibody is sold commercially for $1030/mg. For protein purification we request $3600 which reflects the cost of two anti-HA affinity columns, two heparin sepharose columns, protein G for anti-HA purification, and a mono-S column. The BIAcore instrument is the most quantitative method for characterizing the kinetics of binding of these proteins which gives us insight into their mechanism of binding. We request $5200 for 30 BIAcore CM5 sensor chips, surfactant, O-rings, and coupling reagents. For microbiological supplies we request $7200 which includes media, agar, electroporation cells, amino acids, antibiotics, IPTG, buffers, and miscellaneous chemicals. This reflects the cloning intensive nature of the project.

Travel: A travel allowance is requested for the P.I. to attend one scientific conference focused on transcription or the molecular biology of HIV.

Other expenses: Other expenses are requested to cover tuition for the graduate student, artwork, photography, publication costs, reprints, journals, shipping materials to collaborators, and glass washing.

Personnel costs escalated at 4% per year. Equipment to be purchased during the first year only. Other direct costs are escalated at 6% per year.

BIOGRAPHICAL SKETCH

Give the following information for the key personnel and consultants and collaborators. Begin with the principal investigator/program director. Photocopy this page for each person.

NAME	POSITION TITLE
Carlos F. Barbas, III	Assistant Member

EDUCATION *(Begin with baccalaureate or other initial professional education, such as nursing, and include postdoctoral training.)*

INSTITUTION AND LOCATION	DEGREE	YEAR CONFERRED	FIELD OF STUDY
Eckerd College, St. Petersburg, FL	B.S.	1985	Chemistry & Physics
Texas A&M Univ., College Station, TX	Ph.D.	1989	Organic Chemistry

RESEARCH AND PROFESSIONAL EXPERIENCE: Concluding with present position, list, in chronological order, previous employment, experience, and honors. Key personnel include the principal investigator and any other individuals who participate in the scientific development or execution of the project. Key personnel typically will include all individuals with doctoral or other professional degrees, but in some projects will include individuals at the masters or baccalaureate level provided they contribute in a substantive way to the scientific development or execution of the project. Include present membership on any Federal Government public advisory committee. List, in chronological order, the titles, all authors, and complete references to all publications during the past three years and to representative earlier publications pertinent to this application. If the list of publications in the last three years exceeds two pages, select the most pertinent publications. DO NOT EXCEED TWO PAGES.

Professional Experience:

2/88-9/88 Visiting Scientist, Dept. of Biology, Massachusetts Institute of Technology, Cambridge, MA "Cloning and overproduction of microbial aldolases".

10/89-8/91 Postdoctoral Fellow, Research Institute and Scripps Clinic and the Pennsylvania State University. "Development of antibody cloning/selection strategies and the synthesis of transition-state analogs for the production of catalytic antibodies".

9/91-present Assistant Member, Dept. of Molecular Biology, The Scripps Research Institute, La Jolla, CA

10/92-present Director, Cold Spring Harbor Laboratory Course on Monoclonal Antibodies from Combinatorial Libraries.

Awards or Fellowship:

Investigator Award, Cancer Research Institute, 1993-1997
Scholar of The American Foundation for AIDS Research, 1992-1995
National Institutes of Health Postdoctoral Fellowship, 1990-1991
National Science Foundation Graduate Fellowship, 1986-1989

Publications : (selected from a list of 61)

1. Barbas III, C.F., Kang, A.S., Lerner, R.A. and Benkovic, S.J. (1991) Assembly of combinatorial antibody libraries on phage surfaces: The gene III site. *Proc. Natl. Acad. Sci. USA* **88:**7978-7982.
2. Burton, D.R., Barbas III, C.F., Persson, M.A.A., Koenig, S., Chanock, R.M. and Lerner, R.A. (1991) A large array of human monoclonal antibodies to type 1 human immunodeficiency virus from combinatorial libraries of asymptomatic seropositive individuals. *Proc. Natl. Acad. Sci. USA* **88:**10134-10137.
3. Duchosal, M.A., Eming, S.A., Fischer, P., Leturcq, D., Barbas III, C.F., McConahey, P.J., Caothien, R.H., Thornton, G.B., Dixon, F.J. and Burton, D.R. (1992) Immunization of hu-PBL-SCID mice and the rescue of human monoclonal Fab fragments through combinatorial libraries. *Nature* **355:**358-362.
4. Zebedee, S.L., Barbas III, C.F., Hom, Y.-L., Caothien, R.H., Pa Polla, R., Burton, D.R. and Thornton, G.B. (1992) Human combinatorial antibody libraries to hepatitis B surface antigen. *Proc. Natl. Acad. Sci. USA* **89:**3175-3179.
5. Gram, H., Marconi, L.-A., Barbas III, C.F., Collet, T.A., Lerner, R.A. and Kang, A.S. (1992) *In vitro* selection and affinity maturation of antibodies from a naive combinatorial immunoglobulin library. *Proc. Natl. Acad. Sci. USA* **89:**3576-3580.
6. Barbas III, C.F., Bain, J.D., Hoekstra, D.M. and Lerner, R.A. (1992) Semi-synthetic combinatorial antibody libraries: A chemical solution to the diversity problem. *Proc. Natl. Acad. Sci. USA* **89:**4457-4461.
7. Barbas III, C.F., Bjorling, E., Chiodi, F., Dunlop, N., Cababa, D., Jones, T.M., Zebedee, S.L., Persson, M.A.A., Nara, P.L., Norrby, E. and Burton, D.R. (1992) Recombinant Human Fab fragments neutralize human type 1 immunodeficiency virus *in vitro*. *Proc. Natl. Acad. Sci. USA* **89:**9339-9343.
8. Barbas III, C.F., Crowe, J.E., Cababa, D., Jones, T.M., Zebedee, S.L., Murphy, B.R., Chanock, R.M. and Burton, D.R. (1992) Human monoclonal Fab fragments derived from a combinatorial library bind to respiratory syncytial virus F glycoprotein and neutralize infectivity. *Proc. Natl. Acad. Sci. USA* **89:**10164-10168.
9. Lerner, R.A., Kang, A.S., Bain, J.D., Burton, D.R. and Barbas III, C.F. (1992) Antibodies without immunization. *Science* **258:**1313-1314.
10. Burton, D.R. and Barbas III, C.F. (1993) Human antibodies to HIV-1 by recombinant DNA methods in Immunochemistry of AIDS, ed. Norrby, E. (Karger, Basel, Switzerland), **56:**112-126.

11. Pannekoek, H., vanMeijer, M., Schleef, R.R., Loskutoff, D.J. and Barbas III, C.F. (1993) Functional display of human plasminogen activator inhibitor 1 (PAI-1) on phagemids: Novel perspectives for structure-function analysis by error-prone DNA synthesis. *Gene* **128:**135-140.
12. Williamson, R.A., Burioni, R., Sanna, P.-P., Partridge, L.J., Barbas III, C.F. and Burton, D.R. (1993) Human monoclonal antibodies against a plethora of viral pathogens from single combinatorial libraries. *Proc. Natl. Acad. Sci. USA* **90:**4141-4145.
13. Barbas III, C.F., Collet, T.A., Roben, P., Binley, J., Amberg, W., Hoekstra, D., Cababa, D., Jones, T.M., Williamson, R.A., Pilkington, G.R., Haigwoods, N.L., Satterthwait, A.C., Sanz, I. and Burton, D.R. (1993) Molecular profile of an antibody response to HIV-1 as probed by combinatorial libraries. *J. Mol. Biol.* **230:**812-823.
14. Barbas III, C.F., Rosenblum, J.S. and Lerner, R.A. (1993) Direct selection of antibodies which coordinate metals from semisynthetic combinatorial libraries. *Proc. Natl. Acad. Sci. USA* **14:**6385-6389.
15. Barbas III, C.F., Languino, L.R. and Smith, J.W. (1993) High Affinity Self-Reactive Human Antibodies by Design and Selection: Targeting the Integrin Ligand Binding Site. *Proc. Natl. Acad. Sci. USA* **90:**10003-10007.
16. Barbas III, C.F., Amberg, W., Somincsits, A., Jones, T.M. and Lerner, R.A. (1993) Selection of human anti-hapten antibodies from semisynthetic libraries. *Gene* **137:**57-62.
17. Barbas III, C.F. (1993) Recent advances in phage display. *Curr. Opinion in Biotech.* **4:**526-530.
18. Ditzel, H.J., Barbas, S.M., Barbas III, C.F. and Burton, D.R. (1994) The Nature of the Autoimmune Antibody Repertoire in HIV-1 Infection. *Proc. Natl. Acad. Sci. USA* **91:**3710-3714.
19. Janda, K.D., Lo, C.-H.L., Li, T., Barbas III, C.F., Wirsching, P. and Lerner, R.A. (1994) Direct Selection for a Catalytic Mechanism from Combinatorial Antibody Libraries. *Proc. Natl. Acad. Sci. USA* **91:**2532-2536.
20. Crowe Jr., J.E., Murphy, B.R., Chanock, R.M., Williamson, R.A., Barbas III, C.F. and Burton, D.R. (1994) Recombinant human respiratory syncytial virus (RSV) monoclonal antibody Fab is effective therapeutically when introduced directly into the lungs of RSV-infected mice. *Proc. Natl. Acad. Sci. USA* **91:**2133-2137.
21. Barbas, S.M., Ghazal, P., Barbas III, C.F. and Burton, D.R. (1994) Recognition of DNA by synthetic antibodies. *J. Amer. Chem. Soc.* **116:**2161-2162.
22. Roben, P., Moore, J.P., Thali, M., Sodroski, J., Barbas III, C.F. and Burton, D.R. (1994) Recognition Properties of a Panel of Human Recombinant Fab Fragments to the CD4 Binding Site of GP120 Showing Differing Ability to Neutralise Human Immunodeficiency Virus Type 1. *J. of Virology* **68:**4821-4828.
23. Smith, J.W., Hu, D., Satterthwait, A.C., Pinz-Sweeney, S. and Barbas III, C.F. (1994) Building Synthetic Antibodies as Adhesive Ligands for Integrins. *J. Biol. Chem.* **269:**32788-32795.
24. Rosenblum, J.S. and Barbas III, C.F. (1994) Synthetic Antibodies. *Antibody Engineering*, Oxford Univ. Press (New York, NY) pgs. 89-116.
25. Barbas III, C.F. (1994) The Combinatorial Approach to Human Antibodies. *The Pharmacology of Monoclonal Antibodies: Handbook of Experimental Pharmacology* **113:**243-266.
26. Barbas III, C.F., Hu, D., Dunlop, N., Sawyer, L., Cababa, D., Hendry, R.M., Nara, P.L. and Burton, D.R. (1994) *In vitro* evolution of a neutralizing human antibody to HIV-1 to enhance affinity and broaden strain cross-reactivity. *Proc. Natl. Acad. Sci. USA* **91:**3809-3813.
27. Burton, D.R. and Barbas III, C.F. (1994) Human Antibodies from Combinatorial Libraries. *Advances in Immunology* **57:**191-280.
28. Moore, J.P., Cao, Y., Qing, L., Sattentau, Q.J., Pyati, J., Koduri, R., Robinson, J., Barbas III, C.F., Burton, D.R. and Ho, D.D. (1994) Primary Isolates of Human Immunodeficiency Virus Type 1 are Relatively Resistant to Neutralization by Monoclonal Antibodies to gp120. *J. of Virol.* **69:**101-109.
29. Barbas, S.M. and Barbas III, C.F. (1994) Filamentous Phage Display. *Fibrinolysis* **8:**245-252.
30. Burton, D.R., Pyati, J., Koduri, R., Sharp, S.J., Thornton, G.B., Parren, P.W.H.I., Sawyer, L.S.W., Hendry, M.R., Dunlop, N., Nara, P.L., Lamacchia, M., Garratty, E., Stiehm, E.R., Bryson, Y.J., Cao, Y., Moore, J.P., Ho, D.D. and Barbas III, C.F. (1994) Efficient Neutralization of Primary Isolates of HIV-1 by a Recombinant Human Monoclonal Antibody. *Science* **266:**1024-1026.
31. Barbas III, C.F. and Wagner, J. (1995) Synthetic Human Antibodies: Selecting and Evolving Functional Proteins. *Immunomethods*, in press.
32. Ditzel, H.J., Binley, J.M., Moore, J.P., Sodroski, J., Sawyer, L.S.W., Hendry, R.M., Yang, W.-P., Barbas III, C.F. and Burton, D.R. (1995) Neutralizing Recombinant Human Antibodies to a Conformational V2- and CD4- Binding Site-Sensitive Epitope of HIV-1 gp120 Isolated by Using an Epitope-Masking Procedure. *J. of Immunol.* **154:**893-906.
33. Wu, H., Yang, W.-P. and Barbas III, C.F. (1995) Building Zinc Fingers by Selection: Towards a therapeutic application. *Proc. Natl. Acad. Sci. USA* **92:**344-348.
34. Yang, W.-P., Wu, H. and Barbas III, C.F. (1995) Surface Plasmon Resonance based Kinetic Studies of Zinc Finger-DNA Interactions. *Methods in Immunology*, in press.
35. Wagner, J., Lerner, R.A. and Barbas III, C.F. (1995) Asymmetric Synthesis of Five Haptens Designed for In Vitro Evolution of Antibodies with Peptidase Activity. *J. Med. Chem.*, submitted.
36. Parren, P.W.H.I., Ditzel, H.J., Gulizia, R.J., Binley, J.M., Barbas III, C.F., Burton, D.R. and Mosier, D.E. (1995) Protection against HIV-1 infection in hu-PBL-SCID mice by passive immunization with a neutralizing human monoclonal antibody against the gp120 CD4-binding site. *AIDS*, in press.
37. Yang, W.-P., Green, K., Pinz-Sweeney, S., Briones, A.T. and Barbas III, C.F. (1995) CDR Walking Mutagenesis for the Affinity Maturation of a Potent Human anti-HIV-1 Antibody into the Picomolar Range. *J. Mol. Biol.*, submitted.

BIOGRAPHICAL SKETCH

Give the following information for the key personnel and consultants and collaborators. Begin with the principal investigator/program director. Photocopy this page for each person.

NAME	POSITION TITLE
Herren Wu	Research Associate

EDUCATION *(Begin with baccalaureate or other initial professional education, such as nursing, and include postdoctoral training.)*

INSTITUTION AND LOCATION	DEGREE	YEAR CONFERRED	FIELD OF STUDY
National Taiwan Univ., Taiwan, R.O.C.	B.S.	1986	Chemistry
Univ. of Massachusetts, Amherst, MA	Ph.D.	1993	Mol/Cellular Biol.

RESEARCH AND PROFESSIONAL EXPERIENCE: Concluding with present position, list, in chronological order, previous employment, experience, and honors. Key personnel include the principal investigator and any other individuals who participate in the scientific development or execution of the project. Key personnel typically will include all individuals with doctoral or other professional degrees, but in some projects will include individuals at the masters or baccalaureate level provided they contribute in a substantive way to the scientific development or execution of the project. Include present membership on any Federal Government public advisory committee. List, in chronological order, the titles, all authors, and complete references to all publications during the past three years and to representative earlier publications pertinent to this application. If the list of publications in the last three years exceeds two pages, select the most pertinent publications. DO NOT EXCEED TWO PAGES.

Honors:

1983 Fullen Scholarship, National Taiwan University
1985 Chang-Hsing Chemical Inc. Scholarship, National Taiwan University

Teaching Experience:

1982-1986 Tutor: General Chem., Org. Chem., Physical Chem., Calculus,Math. and Gen. Physics, National Taiwan Univ., China
1988-1989 Teaching Assistant: Biochemistry Laboratory, University of Massachusetts, Amherst, MA (USA)

Research Experience:

1985-1986 Independent Study: Synthesis, structural and functional analysis of a new monomer: 1,4-bis(p-aminocinamoyl) benzene. Dept. of Chemistry, National Taiwan University, Taipei, Taiwan, R.O.C.
1988-1989 Graduate Research Rotation: Overproduction of Escherichia coli ribosomal protein S8 and Production protein in S8 and Production of polyclonal and monoclonal antibodies against guinea pig IgA. Univ. of Mass., Amherst, MA
1989-1993 Ph.D. - Thesis: "Structural and functional aspects of ribosomal protein S8-RNA interactions in *E. coli.* Dept. of Molecular & Cellular Biology, University of Massachusetts, Amherst, MA
1993-present Postdoctoral training: Protein Engineering of DNA-binding proteins. Dept. of Molecular Biology, The Scripps Research Institute, La Jolla, CA 92037

Technical Experience:

Biochemistry: Protein purification and functional analysis: Cell culture/salt precipitation/ion exchange and gel filtration column chromatography/reverse-phase, HPLC/SDS-PAGE/Preparative gel electrophoresis and electroelution/kinetic analysis of protein function.
RNA: *In vitro* transcription/chemical synthesis of RNA/end-labeling/reverse phase and ion exchange HPLC/gel purification/nitrocellulose filter-binding assay.
Biophysics: Protein structure: Circular dichroism(CD)/urea & heat denaturation/nuclear magnetic resonance spectroscopy (NMR).
Genetics: Site-directed mutagenesis/DNA sequencing/DNA recomb. techniques.
Immunology: Tissue culture/purification of immunoglobulin (IgG, IgA, IgM) enzyme linked immunoassays (ELISA) western blotting/immunoelectrophoresis/animal handling/production of polyclonal antibodies/making hybridomas.

Publications:

1. Wu, H., Yang, W.-P., and Barbas III, C.F. (1995). Building zinc fingers by selection: towards a therapeutic application. *Proc. Natl. Acad. Sci. USA* **92**:344-348.
2. Yang, W.-P., Wu, H., and Barbas III, C.F. (1995). Surface plasmon resonance based kinetic studies of zinc finger-DNA interactions. *J. Immun. Methods,* in press.

Publications:

3. Wu, H., Jiang, L., and Zimmermann, R.A. (1994). The binding site for ribosomal protein S8 in 16S rRNA and spc mRNA from *Escherichia coli*: minimum structural requirements and the effects of single bulged bases on S8-RNA interaction. *Nucleic Acids Res.* **22:**1687-1695.
4. Wu, H., Wower, I., and Zimmerman,n R.A. (1993). Mutagenesis of ribosomal protein S8 from *Escherichia coli:* expression, stability, and RNA-binding properties of S8 mutants. *Biochemistry* **32:**4761-4768.
5. Wu, H., Wower, I., and Zimmermann, R.A. Structural and functional relationships of ribosomal protein S8 with 16S rRNA. Manuscript in preparation.

Abstracts:

1. Wower, I., Wu, H., and Zimmermann, R.A. Mutants of ribosomal protein S8 from *E. coli* that are impaired in spc operon regulation and RNA binding. International Conference on the Synthesis of Ribosomes, Cold Spring Harbor, New York, September, 1991.
2. Wu, H., Wower, I., and Zimmermann, R.A. Structural and functional aspects of S8-RNA interactions in *E. coli.* International Conference on the Translational Apparatus, Berlin, Germany, October, 1992.

BIOGRAPHICAL SKETCH

Give the following information for the key personnel and consultants and collaborators. Begin with the principal investigator/program director. Photocopy this page for each person.

NAME	POSITION TITLE
Qiang Liu	Research Associate

EDUCATION *(Begin with baccalaureate or other initial professional education, such as nursing, and include postdoctoral training.)*

INSTITUTION AND LOCATION	DEGREE	YEAR CONFERRED	FIELD OF STUDY
Shanghai Medical Univ., Shanghai, PR China	B.S.	1983	Pharmacy
Shanghai Inst. Bio., Academia Sinica, Shanghai	M.S.	1987	Mol. Biology
Duke University, Durham, NC	Ph.D.	1995	Microbiol/Immuno.

RESEARCH AND PROFESSIONAL EXPERIENCE: Concluding with present position, list, in chronological order, previous employment, experience, and honors. Key personnel include the principal investigator and any other individuals who participate in the scientific development or execution of the project. Key personnel typically will include all individuals with doctoral or other professional degrees, but in some projects will include individuals at the masters or baccalaureate level provided they contribute in a substantive way to the scientific development or execution of the project. Include present membership on any Federal Government public advisory committee. List, in chronological order, the titles, all authors, and complete references to all publications during the past three years and to representative earlier publications pertinent to this application. If the list of publications in the last three years exceeds two pages, select the most pertinent publications. DO NOT EXCEED TWO PAGES.

Professional Experience:

3/95-present	Research Associate, Department of Molecular Biology, The Scripps Research Institute, La Jolla, CA
8/87-8/89	Research Assistant, Lab of Human Genetic Disease, Shanghai Children's Hospital, Shanghai, P.R. China
7/83-9/84	Pharmaceutical Analyst, JiangXi Institute of Pharmaceutical Inspection, NanChang, JiangXi Province, P.R. China

Publications:

1. Liu, Q., and LInney, E. (1993). The Mouse Retinoid-X Receptor-γ Gene: Genomic Organization and Evidence for Functional Isoforms. *Mol. Endocrinol.* **7**:651-658.
2. Hoopes, C.W., Taketo, M., Ozato, K., Liu, Q., Howard, T.A., Linney, E., and Seldin, M.F. (1992). Mapping of the Mouse Rxr Loci Encoding Nuclear Retinoid X Receptors RXRα, RXRβ, and RXRγ. *Genomics* **14**:611-617.
3. Liu, Q., and Linney, E. The Expression of mRXRγ1 is Down Regulated by Retinoic Acids in P19 EC Cell, in preparation.

BIOGRAPHICAL SKETCH

Give the following information for the key personnel and consultants and collaborators. Begin with the principal investigator/program director. Photocopy this page for each person.

NAME	POSITION TITLE
Frederic D. Bushman	Assistant Professor

EDUCATION *(Begin with baccalaureate or other initial professional education, such as nursing, and include postdoctoral training.)*

INSTITUTION AND LOCATION	DEGREE	YEAR CONFERRED	FIELD OF STUDY
Amherst College, Amherst, MA	B.A.	1980	Biology & English
Harvard University, Boston, MA	Ph.D.	1988	Cell & Develop. Biology
Harvard University, Boston, MA	---	---	Ph.D. Research in Molecular Biology
National Institutes of Health, Bethesda, MD	---	---	Ph.D. Research on HIV

RESEARCH AND PROFESSIONAL EXPERIENCE: Concluding with present position, list, in chronological order, previous employment, experience, and honors. Key personnel include the principal investigator and any other individuals who participate in the scientific development or execution of the project. Key personnel typically will include all individuals with doctoral or other professional degrees, but in some projects will include individuals at the masters or baccalaureate level provided they contribute in a substantive way to the scientific development or execution of the project. Include present membership on any Federal Government public advisory committee. List, in chronological order, the titles, all authors, and complete references to all publications during the past three years and to representative earlier publications pertinent to this application. DO NOT EXCEED TWO PAGES.

RESEARCH EXPERIENCE

1993-present Assistant Professor, Infectious Disease Laboratory, The Salk Institute for Biological Studies, La Jolla, CA.

1989-1992 Postdoctoral Fellow, National Institutes of Health, Dr. Kiyoshi Mizuuchi's Laboratory, Bethesda, MD.

1988-1989 Postdoctoral Fellow, Harvard University, Dr. Mark Ptashne's Laboratory, Boston, MA.

1982-1988 Graduate Student, Harvard University, Dr. Mark Ptashne's Laboratory, Boston, MA.

1980-1982 Research Assistant, Worcester Foundation for Experimental Biology, Dr. W.R. Crain's Laboratory, Worcester, MA.

APPOINTMENTS AND AWARDS

1995-1997 Editorial Board, Journal of Virology

1992 Special Fellowship, Leukemia Society of America, Inc.

1989-1992 Fellowship, Leukemia Society of America, Inc.

PATENT APPLICATIONS

1990 "In vitro retroviral integration assay." US patent application No. 07/472,186.

1994 "Method for site-specific integration of nucleic acids and related products" US patent application submitted

TEACHING EXPERIENCE

1993, 1995 Instructor for the graduate course "Human Retroviruses" (Other instructors include: Drs. Flossie Wong-Staal and Thomas Hope), University of California, San Diego, CA.

1986-1987 Instructed first year graduate students (3) during rotations in the laboratory of Dr. Mark Ptashne, Harvard University, Boston, MA.

1984 Teaching Fellow for "Gene Regulation" (Instructor: Dr. Mark Ptashne), Harvard University, Boston, MA.

1983 Teaching Fellow for "Introductory Biology" (Instructors: Dr. Lawrence Bogorad, Dr. Woodward Hastings, and Dr. John Kimball), Harvard University, Boston, MA.

PUBLICATIONS

1. W.R. Crain, D.S., Durica, A.D. Cooper, K. Van Doren and F.D. Bushman. (1982) Structure and developmental expression of actin genes in the sea urchin. *In:* "Muscle Development: Molecular and Cellular Control" (M. Pearson and H.F. Epstein, eds.) Academic Press, London, pp. 177-191.

2. F.D. Bushman and W.R. Crain. (1983) Conserved pattern of embryonic actin gene expression in several sea urchins and a sand dollar. **Dev. Biol.** 98: 429-436.
3. W.R. Crain and F.D. Bushman. (1983) Transcripts of paternal and maternal actin gene alleles are present in interspecific sea urchin embryo hybrids. **Dev. Biol.** 100: 190-196.
4. F.D. Bushman, J.E., Anderson, S.C. Harrison and M. Ptashne. (1985) Ethylation interference and X-ray crystallography identify similar interactions between 434 repressor and operator. **Nature** 316: 651-653.
5. F.D. Bushman and M. Ptashne. (1986) Activation of transcription by the bacteriophage 434 repressor. **Proc. Natl. Acad. Sci. USA** 83: 9353-9357.
6. F.D. Bushman and M. Ptashne. (1988) Turning lambda Cro into a transcriptional activator. **Cell** 58: 1163-1171.
7. F.D. Bushman, C. Shang and M. Ptashne. (1989) A single glutamic acid residue plays a key role in the transcriptional activation function of lambda repressor. **Cell** 58: 1163-1171.
8. R. Craigie, T. Fujiwara and F.D. Bushman. (1990) The in protein in Moloney murine leukemia virus processes the viral DNA ends and accomplishes their integration in vitro. **Cell** 62: 829-837.
9. F.D. Bushman, T. Fujiwara and R. Craigie. (1990) Retroviral DNA integration directed by HIV IN protein in vitro. **Science** 249: 1555-1558.
10. F.D. Bushman and R. Craigie. (1990) Sequences at the ends of Moloney murine leukemia virus DNA required for integration in vitro. **J. Virol.** 64: 5645-5648.
11. F.D. Bushman and R. Craigie. (1991) Activities of human immunodeficiency virus (HIV) integration protein in vitro: Specific cleavage and integration of HIV DNA. **Proc. Natl. Acad. Sci. USA** 88: 1339-1343.
12. R. Craigie, K. Mizuuchi, F.D. Bushman and A. Engelman. (1991) A rapid in vitro assay for HIV DNA integration. **Nuc. Acids Res.** 19: 2729-2734.
13. F.D. Bushman. (1992) Activators, deactivators, and deactived activators. **Current Biology** 2: 673-657.
14. F.D. Bushman and R. Craigie. (1992) Integration of human immunodeficiency virus DNA: adduct interference analysis of required sites. **Proc. Natl. Acad. Sci. USA** 89: 3458-3462.
15. F.D. Bushman. (1993) The bacteriophage 434 right operator: Roles of O_R1, O_R2, and O_R3. **J. Mol. Biol.** 230: 28-40.
16. F.D. Bushman, A. Engelman, I. Palmer, P. Wingfield and R. Craigie. (1993) Domains of the integrase protein of human immunodeficiency virus type 1 responsible for polynucleotidyl transfer and zinc binding. **Proc. Natl. Acad. Sci. USA** 90: 3428-3432.
17. A. Engelman, F.D. Bushman and R. Craigie. (1993) Identification of discrete functional domains of HIV-1 integrase and their organization within an active multimeric complex. **EMBO J.** 12: 3269-3276.
18. F.D. Bushman. (1993) Dodging the genes. **Current Biology** 3: 533-535.
19. F.D. Bushman and B. Wang. (1994) Rous Sarcoma Virus Integrase Protein: Mapping Functions for Catalysis and Substrate Binding. **J. Virol.** 68:2215-2223.
20. D. Pruss, F.D. Bushman and A.P. Wolffe. (1994) HIV integrase directs integration to sites of severe DNA distortion within the nucleosome core. **Proc. Natl. Acad. Sci. USA** 91: 5913-5917.
21. F.D. Bushman. (1994) Tethering human immunodeficiency virus 1 integrase to a DNA site directs integration to nearby sequences. **Proc. Natl. Acad. Sci. USA** 91: 9233-9237
22. D. Pruss, R. Reeves, F.D. Bushman and A.P. Wolffe. (1994) The influence of DNA and nucleosome structure on integration events directed by HIV integrase. **J. Biol. Chem.** 269: 25031-25041.
23. F.D. Bushman. (1995) Targeting retroviral integration. **Science** 267: 1443-1444.
24. M.D. Miller, B. Wang, and F.D. Bushman. (1995) Integration of incomplete HIV-1 reverse transcription products. **J. Virol.** (in press).
25. M.D. Miller and F.D. Bushman. Ini1 for integration. **Current Biology** (in press).

BIOGRAPHICAL SKETCH

Give the following information for the key personnel and consultants and collaborators. Begin with the principal investigator/program director. Photocopy this page for each person.

NAME	POSITION TITLE
Moore, John P., Ph.D.	Staff Investigator

EDUCATION *(Begin with baccalaureate or other initial professional education, such as nursing, and include postdoctoral training.)*

INSTITUTION AND LOCATION	DEGREE	YEAR CONFERRED	FIELD OF STUDY
Downing College, Cambridge University, UK	BA	1978	Biochemistry
	M. Phil.	1979	Biochemistry
	M.A.	1982	Biochemistry
	Ph.D.	1982	Biochemistry

RESEARCH AND PROFESSIONAL EXPERIENCE: Concluding with present position, list, in chronological order, previous employment, experience, and honors. Key personnel include the principal investigator and any other individuals who participate in the scientific development or execution of the project. Key personnel typically will include all individuals with doctoral or other professional degrees, but in some projects will include individuals at the masters or baccalaureate level provided they contribute in a substantive way to the scientific development or execution of the project. Include present membership on any Federal Government public advisory committee. List, in chronological order, the titles, all authors, and complete references to all publications during the past three years and to representative earlier publications pertinent to this application. If the list of publications in the last three years exceeds two pages, s lect the most pertinent publications. DO NOT EXCEED TWO PAGES.

POSITIONS HELD

1982-1986	Postdoctoral Research Assistant, Dept. of Biochemistry, Cambridge University.
1986-1987	Visiting Scientist, Ludwig Institute for Cancer Research (Dr. G.I. Evan's lab), Cambridge.
1987-1988	Senior Research Fellow, Dept. of Veterinary Pathology (Dr. W. Jarrett's lab), University of Glasgow.
1989-1991	Research Fellow, Institute of Cancer Research (Dr. R. Weiss' lab), Chester Beatty Laboratories, London.
1992-	Staff Investigator,The Aaron Diamond AIDS Research Center, New York.
1992-	Associate Professor of Microbiology, New York University School of Medicine.

SELECTED PUBLICATIONS

Moore JP, Jarrett RF. Sensitive ELISA for gp120 and gp160 surface glycoproteins of HIV-1. AIDS Res. Human Retroviruses 1988; 4:369-379

Moore JP, Wallace LA, Follett EAC, McKeating JA. An ELISA for antibodies to the envelope glyco-protein of divergent strains of HIV-1. AIDS 1989; 3:155-163.

Moore JP. Simple methods for monitoring HIV-1 and HIV-2 gp120 binding to sCD4 by ELISA: HIV-2 has a 25-fold lower affinity than HIV-1 for sCD4. AIDS 1990; 4:297-305.

Moore JP, McKeating JA, Jones IM, Stephens PE, Clements G, Thomson S, Weiss RA. Characterization of recombinant gp120 and gp160 from HIV-1: binding to monoclonal antibodies and sCD4. AIDS 1990; 4:307-315.

Moore JP, McKeating RA, Weiss RA, Sattentau QJ. Dissociation of gp120 from HIV-1 virions induced by soluble CD4. Science 1990; 250:1139-1142.

Clements GJ, Price-Jones MJ, Stephens PE, Sutton C, Schulz TF, Clapham PR, McKeating JA, McClure MO, Thomson S, Marsh M, Kay J, Weiss RA, **Moore JP**. The V3 loops of the HIV-1 and HIV-2 surface glycoproteins contain proteolytic cleavage sites: a possible function in viral fusion? AIDS Res. Human Retroviruses 1991; 7:3-16.

McKeating JA, McKnight A, **Moore JP**. Differential loss of envelope glycoprotein gp120 from virions of human immunodeficiency virus type 1 isolates: effect on infectivity and neutralization. J. Virol. 1991; 65:852-860.

Moore JP, McKeating JA, Norton WA, Sattentau QJ. Direct measurement of soluble CD4 binding to human immunodeficiency virus type 1 virions: gp120 dissociation and its implications for the virus-cell binding and fusion reactions and their neutralization by soluble CD4. J. Virol. 1991; 65:1133-1140.

Moore JP, Morikawa Y, Jones IM. Binding of recombinant HIV-1 and HIV-2 SU glycoproteins to sCD4. J. Acq. Immun. Def. Synd. 1991; 4:442-443.

Sattentau QJ, **Moore JP**. Conformational changes induced in the human immunodeficiency virus envelope glycoproteins by soluble CD4 binding. J. Exp. Med. 1991; 174:406-415.

Cordell J, **Moore JP**, Dean CJ, Klasse PJ, Weiss RA, McKeating JA. Rat monoclonal antibodies to nonoverlapping epitopes of human immunodeficiency virus type 1 gp120 block CD4 binding in vitro. Virology 1991; 185:72-79.

Moore JP, Nara PL. The role of the V3 domain of gp120 in HIV infection. AIDS 1991; 5(2):S21-S33.

Moore JP, McKeating JA, Huang Y, Ashkenazi A, Ho DD. Virions of primary human immunodefi-ciency virus type 1 isolates resistant to soluble CD4 (sCD4) neutralization differ in sCD4 binding and glycoprotein gp120 retention from sCD4-sensitive isolates. J. Virol. 1992; 66:235-243.

Moore JP, Klasse PJ. Thermodynamic and kinetic analysis of sCD4 binding to HIV-1 virions and of gp120 dissociation. AIDS Res. Human Retroviruses. 1992;8:443-450.

McKeating JA, **Moore JP**, Ferguson M, Marsden HS, Graham S, Almond JW, Evans D, Weiss RA. Monoclonal antibodies to the C4 region of human immunodeficiency virus type 1 gp120: use in topological analysis of a CD4 binding site. AIDS 1992; 8:451-459.

Moore JP, Simpson G, McKeating JA, Burakoff SJ, Schreiber SL, Weiss RA. CPF-DD is a inhibitor of infection by human immunodeficiency virus and other enveloped viruses in vitro. Virology. 1992; 188:537-544.

Moore JP, Sattentau QJ, Klasse PJ, Burkly L. A monoclonal antibody to CD4 domain 2 blocks soluble CD4-induced conformational changes in the envelope glycoproteins of human immunodeficiency virus type 1 (HIV-1) and HIV-1 infection of CD4+ cells. J. Virol. 1992; 66:4784-4793.

Schultz TF, Jameson BA, Lopalco L, Siccardi AG, Weiss RA, **Moore JP**. Conserved structural features in the interaction between retroviral surface and transmembrane glycoproteins? AIDS Res. Human Retroviruses 1992; 8:1585-1594.

Moore JP, Jameson BA, Weiss RA, Sattentau QJ. HIV-cell fusion. In "Viral Fusion Mechanism." J. Bentz, ed.) CRC Press Inc. Boca Raton, USA pp 233-289.

Moore JP, Ho DD. Antibodies to discontinuous or conformationally sensitive epitopes on the gp120 glycoprotein of human immunodeficiency virus type 1 are highly prevalent in sera of infected humans. J.Virol. 1993;67:863-875.

Moore JP. The reactivities of HIV-1+ human sera with solid-phase V3 loop peptides can be poor predictors of their reactivities with V3 loops on native gp120 molecules. AIDS Res. Human Retroviruses 1993;9:209-219.

Moore JP. A monoclonal antibody to the CDR-3 region of CD4 inhibits soluble CD4 binding to virions of human immunodeficiency virus type 1. J.Virol. 1993;67:3656-3659.

Moore JP, Burkly LC, Connor RI, Cao Y, Tizard R, Ho DD, Fisher RA. Adaptation of two primary, macrophage-tropic human immunodeficiency virus type 1 isolates to growth in transformed T-cell lines correlates with alterations in the responses of their envelope glycoproteins to soluble CD4. AIDS Res. Human Retroviruses 1993;9:529-539.

Thali M, Moore JP, Furman C, Charles M, Ho DD, Robinson J, Sodroski J. Characterization of conserved human immunodeficiency virus type 1 gp120 neutralization epitopes exposed upon gp120-CD4 binding. J. Virol. 1993;67:3978-3988.

Moore JP, Thali M, Jameson BA, Vignaux F, Lewis GK, Poon S-W, Charles M, Fung MS, Sun B, Durda PJ, Åkerblom L, Wahren B, Ho DD, Sattentau QJ, Sodroski J. Immunochemical analysis of the gp120 surface glycoprotein of human immunodeficiency virus type 1: Probing the structure of the C4 and V4 domains and the interaction of the C4 domain with the V3 loop. J. Virol. 1993;67:4785-4796.

Moore JP, Sweet RW. The HIV gp120-CD4 interaction: A target for pharmacological or immunological intervention? Perspectives in Drug Discovery and Design 1993;1:235-250.

Moore JP, Sattentau QJ, Yoshiyama H, Thali M, Charles M, Sullivan N, Poon S-W, Pinkus M, Robey G, Fung MS, Traincard F, Robinson JE, Ho DD, Sodroski J. Probing the structure of the V2 domain of the human immunodeficiency virus type 1 surface glycoprotein gp120 with a panel of eight monoclonal antibodies: The human immune response to the V1 and V2 domains. J.Virol. 1993;67:6136-6151.

Sattentau QJ, **Moore JP**, Vignaux F, Traincard F, Poignard P. Conformational changes induced in the envelope glycoproteins of the human and simian immunodeficiency viruses by soluble receptor binding. J.Virol. 1993;67:7383-7393.

Moore JP, Yoshiyama H, Ho DD, Robinson JE, Sodroski J. Antigenic variation in gp120s from molecular clones of HIV-1 LAI. AIDS Res. Human Retroviruses 1993;9:1179-1187.

Moore JP, Sattentau QJ, Wyatt R, Sodroski J. Mapping the topology of the surface glycoprotein gp120 of human immunodeficiency virus type 1 with a panel of monoclonal antibodies. J.Virol. 1994;68:469-484.

Sattentau QJ, Vignaux F, Sodroski J, **Moore JP**. Exposure and modulation of neutralizing epitopes on the outer envelope glycoprotein of the human immunodeficiency virus type 1. J. Immunol. (submitted).

Moore JP, Cao Y, Conley AJ, Wyatt R, Robinson J, Gorny M, Zolla-Pazner S, Ho DD, Koup RA. Studies with monoclonal antibodies to the V3 region of HIV-1 gp120 reveal limitations to the utility of solid-phase peptide binding assays. JAIDS. (in press).

O'Brien WA, Mao S-H, Cao Y, **Moore JP**. Macrophage- and T-cell-tropic chimeric strains of human immunodeficiency virus type 1 differ in their susceptibility to neutralization by soluble CD4 at different temperatures. J. Virol. (submitted).

BIOGRAPHICAL SKETCH

Give the following information for the key personnel and consultants and collaborators. Begin with the principal investigator/program director. Photocopy this page for each person.

NAME	POSITION TITLE
Flossie Wong-Staal, Ph.D.	Professor of Medicine and Biology

EDUCATION *(Begin with baccalaureate or other initial professional education, such as nursing, and include postdoctoral training.)*

INSTITUTION AND LOCATION	DEGREE	YEAR CONFERRED	FIELD OF STUDY
University of California, Los Angeles, CA	B.A.	1968	Bacteriology (magna cum laude)
University of California, Los Angeles, CA	Ph.D.	1972	Molecular Biology

RESEARCH AND PROFESSIONAL EXPERIENCE: Concluding with present position, list, in chronological order, previous employment, experience, and honors. Key personnel include the principal investigator and any other individuals who participate in the scientific development or execution of the project. Key personnel typically will include all individuals with doctoral or other professional degrees, but in some projects will include individuals at the masters or baccalaureate level provided they contribute in a substantive way to the scientific development or execution of the project. Include present membership on any Federal Government public advisory committee. List, in chronological order, the titles, all authors, and complete references to all publications during the past three years and to representative earlier publications pertinent to this application. If the list of publications in the last three years exceeds two pages, select the most pertinent publications. DO NOT EXCEED TWO PAGES.

Professional Experience

1969-1970 Teaching Assistant, University of California, Los Angeles, California
1970-1972 Research Assistant, University of California, Los Angeles, California
1972-1973 Post-doctoral Fellow, University of California, San Diego, California
1973-1975 Fogarty Fellow, National Cancer Institute, Bethesda, Maryland
1975-1976 Visiting Associate, National Cancer Institute, Bethesda, Maryland
1976-1978 Cancer Expert, National Cancer Institute, Bethesda, Maryland
1978-1981 Senior Investigator, National Cancer Institute, Bethesda, Maryland
1982-1989 Chief, Molecular Genetics of Hematopoietic Cells Section, National Cancer Institute, Bethesda, Maryland
1990-Pres. Florence Seeley Riford Chair in AIDS Research, Professor of Medicine and Biology, University of California, San Diego, La Jolla, California

Societies

American Society for Virology, Charter Member

Honors and Awards

1967-1968 Regents Scholarship, UCLA
1968 Phi Beta Kappa
1968 Regents Fellowship, UCLA
1972 Women Graduates of the Year Award, UCLA
1985 Visiting Professorship, Institute of General Pathology; First University of Rome, Italy
1987 Outstanding Scientific Award, The Chinese Medical and Health Association
1991 The Excellence 2000 Award, United States Pan Asian American Chamber of Commerce and the Organization of Chinese American Women

Editorial Board

1984-Pres. Gene Analysis Techniques
1984-1994 Cancer Letters
1987-Pres. Leukemia

1987	Cancer Research
1987-Pres.	AIDS Research and Human Retroviruses (Section Editor)
1987-Pres.	DNA and Cell Biology (Section Editor)
1987-1990	Microbial Pathogenesis
1987-Pres.	AIDS: An International Journal
1988-Pres.	International Journal of Acquired Immunodeficiency Syndrome
1988-Pres.	Oncogene
1990-Pres.	Journal of Virology

PUBLICATIONS: 15 of 212

1. Shaw, G.M., Hahn, B.H., Arya, S.K., Groopman, J.E., Gallo, R.C., and Wong-Staal, F.: Molecular characterization of human T-cell leukemia (lymphotropic) virus type III in the acquired immune deficiency syndrome. Science 226: 1165-1171, 1984.
2. Ratner, L., Haseltine, W., Patarca, R., Livak, K.J., Starcich, B., Josephs, S.F., Doran, E.R., Rafalski, J.A., Whitehorn, E.A., Baumeister, K., Ivanoff, L., Petteway, S.R., Jr., Pearson, M.L., Lautenberger, J.A., Papas, T.S., Ghrayeb, J., Chang, N.T., Gallo, R.C., and Wong-Staal, F.: Complete nucleotide sequence of the AIDS virus HTLV-III. Nature 313: 277-284, 1985.
3. Fisher, A.G., Collalti, E., Ratner, L., Gallo, R.C., and Wong-Staal, F.: A molecular clone of HTLV-III with biological activity. Nature 316: 262-265, 1985.
4. Arya, S.K., Guo, C., Joseph, S.F., and Wong-Staal, F.: Transactivator gene of human T-lymphotropic virus type III (HTLV-III). Science 229: 69-73, 1985.
5. Wong-Staal, F., Shaw, G.M., Hahn, B.H., Salahuddin, S.Z., Popovic, M., Markham, P., Redfield, R., and Gallo, R.C.: Genomic diversity of human T-lymphotropic virus type III (HTLV-III). Science 229: 759-762, 1985.
6. Fisher, A.G., Ensoli, B., Looney, D., Rose, A., Gallo, R.C., Saag, M.S., Shaw, G.M., Hahn, B.H., and Wong-Staal, F.: Biologically diverse molecular variants within a single HIV-1 isolate. Nature 334: 444-447, 1988.
7. Steffy, K., Kraus, G., Looney, D.J., and Wong-Staal, F.: Role of the fusogenic peptide sequence in syncytium induction and infectivity of human immunodeficiency virus type. J. of Virol. 66: 4532-4535, 1992.
8. Ojwang, J., Hampel, A., Looney, D.J., and Wong-Staal, F., and Rappaport, J.: A hairpin ribozyme inhibits expression of diverse strains of human immunodeficiency virus type 1. Proc. Natl. Acad. Sci., USA, 90: 6340-6344, 1993.
9. Yu, M., Ojwang, J., Yamada, O., Hampel, A., Rappaport, J., Looney, D., and Wong-Staal, F.: A hairpin ribozyme inhibits expression of diverse strains of human immunodeficiency virus type 1. Proc. Natl. Acad. Sci., USA, 90: 6340-6344, 1993.
10. Yamada, O., Yu, M., Yee, J-K., Kraus, G., Looney, D.J., Wong-Staal, F.: Intracellular immunization of human T-cells with a hairpin ribozyme against human immunodeficiency virus type 1. Gene Therapy, 1: 39-45, 1994.
11. Leavitt, M., Yu, M., Yamada, O., Kraus, G., Looney, D., Poeschla, E., and Wong-Staal, F.: Transfer of an anti-HIV-1 ribozyme gene into primary human lymphocytes. Human Gene Therapy 5: 115-1120, 1994.
12. Leavitt, M., Yu, M., Yamada, O., Kraus, G., Looney, D., Poeschla, E., and Wong-Staal, F.: Transfer of an anti-HIV-1 ribozyme gene into primary human lymphocytes. Human Gene Therapy 5: 1115-1120, 1994.
13. Yamada, O., Kraus, G., Leavitt, M.C., Yu, M., and Wong-Staal, F.: Activity and cleavage site specificity of an anti-HIV-1 hairpin ribozyme in human T-cells. Virology, 205: 121-126, 1994 .
14. Yu, M., Leavitt, M., Maruyama, M., Yamada, O., Young, D., Ho, A., and Wong-Staal, F.: Intracellular immunization of human hematopoietic stem cells with a ribozyme against HIV-1. Proc. Natl. Acad. Sci. USA, 92: 699-703, 1995.
15. Yu, M., Poeschla, E., Yamada, O., Degrandis, P., Leavitt, M.C., Heusch, M., Yee, J-K., Wong-Staal, F., and Hampel, A.: In vitro and in vivo characterization of a second functional hairpin ribozyme against HIV-1. Virology, 206: 381-386, 1995.

BIOGRAPHICAL SKETCH

Give the following information for the key personnel and consultants and collaborators. Begin with the principal investigator/program director. Photocopy this page for each person.

NAME	POSITION TITLE
Joel M. Gottesfeld, Ph.D.	**Associate Member**

EDUCATION (*Begin with baccalaureate or other initial professional education, such as nursing, and include postdoctoral training.*)

INSTITUTION AND LOCATION	DEGREE	YEAR CONFERRED	FIELD OF STUDY
University of California, Berkeley, CA	B.A.	1971	Biochemistry
Merton College, Oxford University, England	M.Sc.	1973	Biochemistry
California Institute of Technology, Pasadena,CA	Ph.D.	1975	Biology/Biochem

RESEARCH AND PROFESSIONAL EXPERIENCE: Concluding with present position, list, in chronological order, previous employment, experience, and honors. Key personnel include the principal investigator and any other individuals who participate in the scientific development or execution of the project. Key personnel typically will include all individuals with doctoral or other professional degrees, but in some projects will include individuals at the masters or baccalaureate level provided they contribute in a substantive way to the scientific development or execution of the project. Include present membership on any Federal Government public advisory committee. List, in chronological order, the titles, all authors, and complete references to all publications during the past three years and to representative earlier publications pertinent to this application. DO NOT EXCEED TWO PAGES.

I. PROFESSIONAL EXPERIENCE

1971-1972	Graduate research under Drs. G.K. Radda and I.O. Walker, Department of Biochemistry, Oxford University, England
1972-1975	Graduate study for Ph.D. under the direction of Professor James Bonner, Division of Biology, California Institute of Technology, Pasadena, CA
1975-1978	Postdoctoral research under Dr. J.B. Gurdon, MRC Laboratory of Molecular Biology, Cambridge, England
1978-1983	Assistant Member, Division of Cellular Biology, Research Institute of Scripps Clinic, La Jolla, CA
1983-1984	Associate Member, Division of Cellular Biology, Research Institute of Scripps Clinic, La Jolla, CA
1984-1989	Associate Member, Department of Molecular Biology, Research Institute of Scripps Clinic, La Jolla, CA
1989-1992	Chief and Member, Division of Developmental Biology, Medical Biology Institute, La Jolla, CA
1992-Present	Associate Member, Department of Molecular Biology, The Scripps Research Institute, La Jolla, CA

II. HONORS

1968	University of California Alumni Foundation Scholarship
1971-1972	Fulbright Scholar, Merton College, Oxford University
1975-1978	Helen Hay Whitney Foundation Postdoctoral Fellowship
1983-1986	Biomedical Sciences Study Section, BI-3 NIH
1985	Molecular Cytology Study Section, NIH
1985	American Cancer Society Faculty Research Award
1986	Elected Chairman, 1990 Gordon Research Conference on Chromatin Structure
1988-1989	Molecular Biology Study Section, NIH
1990	Associate Editor, "Gene Expression"

III. REPRESENTATIVE PUBLICATIONS

34. Peck, L.J., L.S. Millstein, P. Eversole-Cire, J.M. Gottesfeld, and A. Varshavsky. Transcriptionally inactive oocyte-type 5S RNA genes of *Xenopus laevis* are complexed with TFIIIA *in vitro*. *Mol. Cell. Biol. 7*:3503-3510, 1987.
35. Millstein, L., P. Eversole-Cire, J. Blanco, and J.M. Gottesfeld. Differential transcription of *Xenopus* oocyte and somatic-type 5S RNA genes in a *Xenopus* oocyte extract. *J. Biol. Chem. 262*:17100-17110, 1987.
37. Blanco, J. and J.M. Gottesfeld. *Xenopus* transcription factor IIIA forms a complex of covalent character with 5SDNA. *Nucl. Acids Res. 16*:11267-11284, 1988.
38. Blanco, J., L. Millstein, S. Dillworth, M. Razik, C. Cote, and J.M. Gottesfeld. Multiple forms of TFIIIA control the differential expression of *Xenopus* oocyte and somatic-type 5S RNA genes. In *UCLA Symp. Molecular and Cellular Biology, New Series: DNA-Protein Interactions in Transcription, 95*:175-186 (J. Gralla, ed.) Alan R. Liss, Inc., New York, 1988.
39. Razik, M.A. and J.M. Gottesfeld. Pathways of nucleoprotein assembly on 5S RNA genes in a *Xenopus* oocyte S-150 extract. *Nucl. Acids Res. 17*:4117, 1989.
40. Schroth, G.P., G.R. Cook, E.M. Bradbury, and J.M. Gottesfeld. Transcription factor IIIA induced bending of the *Xenopus* somatic 5S gene promoter. *Nature 340*:487-488, 1989.
41. Blanco, J., L. Millstein, M.A. Razik, S. Dilworth, C. Cote, and J.M. Gottesfeld. Two TFIIIA activities regulate expression of the *Xenopus* 5S RNA gene families. *Genes and Development 3*:1602-1612, 1989.
42. Millstein, L.S. and J.M. Gottesfeld. Control of gene expression in eukaryotic cells: lessons from class III genes. *Current Opinion in Cell Biology 1*:497-502, 1989.
43. Keller, H.J., Q. You, P. Romaniuk, and J.M. Gottesfeld. Additional intragenic promoter elements of the *Xenopus* 5S RNA genes upstream of the TFIIIA-binding site. *Molecular Cell Biology 10*:5166-5176, 1990.
44. Engelke, D.R. and J.M. Gottesfeld. Chromosomal footprinting of transcriptionally active and inactive oocyte-type 5S RNA genes of *Xenopus laevis*. *Nucl. Acids Res. 18*:6031-6037,1990.
45. Lee, M.S., J.M. Gottesfeld, and P.E. Wright. Zinc is required for folding and binding of a single zinc finger to DNA. *FEBS Lett. 279*:289-294, 1991.
46. Schroth, G.P., J.M. Gottesfeld, and E.M. Bradbury. TFIIIA induced DNA bending: Effect of low ionic strength electrophoresis buffer conditions. *Nucl. Acids Res. 19*:511-516, 1991.
47. Lee, D.K., R.K. Evans, J. Blanco, J.M. Gottesfeld, and J.D. Johnson. Contacts between 5S DNA and *Xenopus* TFIIIA identified using 5-azido-2'-deoxyuridine substituted DNA. *J. Biol. Chem. 266*:16478-16484, 1991.
48. Liao, X-B., K.R. Clemens, L.L. Tennant, P.E. Wright and J.M. Gottesfeld. Specific interaction of the first three fingers of TFIIIA with the internal control region of the *Xenopus* 5S RNA gene. *J. Mol. Biol. 223*:857-871, 1992.
49. Keller, H.J., P.J. Romaniuk and J.M. Gottesfeld. Interaction of *Xenopus* TFIIIC with the TFIIIA-5S RNA gene complex. *J. Biol. Chem. 267*:18190-18198, 1992.
50. Clemens, K.R., X-B Liao, V.J. Wolf, P.E. Wright and J.M. Gottesfeld. Definition of the binding sites of individual zinc fingers in the TFIIIA-5S RNA gene complex. *Proc. Natl. Acad. Sci. USA 89*:10822-10826, 1992.
51. Hartl, P., J.M. Gottesfeld and D. Forbes. Mitotic repression of transcription *in vitro*: Analysis of mechanisms. *J. Cell Biol. 120*:613-624, 1993.
52. Clemens, K.R., V. Wolf, S.J. McBryant, P. Zhang, X-B. Liao, P.E. Wright and J.M. Gottesfeld. Molecular Basis for Specific Recognition of Both RNA and DNA by a Zinc Finger Protein. *Science,260*:530-533, 1993.
53. Gottesfeld, J.M., V.J. Wolf, D.J. Forbes and P. Hartl. Mitotic Repression of RNA Polymerase III Transcription *In Vitro* is Mediated by Phosphorylation of a TFIIIB Component. *Science* 263:81-84, 1994
54. Wolf, V.J., T. Dang, D.J. Forbes and J.M. Gottesfeld. Maturation Promoting Factor (p34cdc2-Cyclin B Kinase) Establishes the Oocyte/Somatic Switch in Xenopus 5S RNA Gene Expression *In Vitro*. *Mol. Cell. Biol.*, In Press.
55. Brown, M.L., G.P. Schroth, J.M. Gottesfeld and D.P. Bazett-Jones. Requirements for the TFIIA Induced Bending of the 5SrRNA Gene Promoter. In preparation.

BIOGRAPHICAL SKETCH

Give the following information for the key personnel and consultants and collaborators. Begin with the principal investigator/program director. Photocopy this page for ach person.

NAME	POSITION TITLE
Joseph G. Sodroski	Associate Professor in Pathology

EDUCATION *(Begin with baccalaureate or other initial professional education, such as nursing, and include postdoctoral training.)*

INSTITUTION AND LOCATION	DEGREE	YEAR CONFERRED	FIELD OF STUDY
Allentown College of St. Francis de Sales	B.S.	1976	Medicine
Jefferson Medical College	M.D.	1980	Medicine

RESEARCH AND PROFESSIONAL EXPERIENCE: Concluding with present position, list, in chronological order, previous employment, experience, and honors. Key personnel include the principal investigator and any other individuals who participate in the scientific development or execution of the project. Key personnel typically will include all individuals with doctoral or other professional degrees, but in some projects will include individuals at the masters or baccalaureate level provided they contribute in a substantive way to the scientific development or execution of the project. Include present membership on any Federal Government public advisory committee. List, in chronological order, the titles, all authors, and complete references to all publications during the past three years and to representative earlier publications pertinent to this application. If the list of publications in the last three years exceeds two pages, select the most pertinent publications. DO NOT EXCEED TWO PAGES.

POSTDOCTORAL TRAINING:

Internships and Residencies:

1980-1981 Intern in Medicine, New England Deaconess Hospital, Boston, MA

Research Fellowships:

1981 Research Fellow in Microbiology, Dana-Farber Cancer Institute, Harvard School of Public Health, Boston, MA

ACADEMIC APPOINTMENTS:

1984 Instructor in Pathology, Dana-Farber Cancer Institute, Harvard Medical School, Boston, MA

1986 Assistant Professor in Pathology, Dana-Farber Cancer Institute, Harvard Medical School, Boston, MA

1990 Associate Professor in Pathology, Dana-Farber Cancer Institute, Harvard Medical School, Boston, MA

SELECTED PUBLICATIONS

1. Sodroski JG, Rosen CA, Haseltine WA. Trans-acting Transcriptional Activation of the Long Terminal Repeat of Human T-Lymphotropic Viruses in Infected Cells. Science. 1984;225:381-385.
2. Sodroski JG, Rosen CR, Wong-Staal F, Popovic M, Arya S, Gallo RC, Haseltine WA. Trans-acting Transcriptional Activation of the Long Terminal Repeat of Human T-Cell Lymphotropic Virus Type III (HTLV-III). Science. 1985;227:171-173.
3. Sodroski JG, Patarca R, Rosen CA, Wong-Staal F, Gallo RC, and Haseltine WA. Location of trans-activating Region on the Genome of HTLV-III/LAV. Science. 1985;229:74-77.
4. Sodroski JG, Rosen CA, Dayton AI, Goh WC, Haseltine WA. Evidence for a Second Post-transcriptional trans-activator Required for HTLV-III/LAV Replication. Nature. 1986;321:412-417.
5. Sodroski JG, Goh WC, Rosen CA, Campbell K, Haseltine WA. Role of the HTLV-III Envelope in Syncytium Formation and Cytopathicity. Nature. 1986;322:470-474.

6. Kowalski M, Potz J, Basiripour L, Dorfman T, Goh WC, Terwilliger E, Dayton A, Rosen C, Haseltine WA, Sodroski J. Functional Regions of the Human Immunodeficiency Virus Envelope Glycoprotein. Science. 1987;237:1351-1355.
7. Hussey R, Richardson N, Kowalski M, Brown N, Hsiu-Ching C, Siliciano R, Dorfman T, Walker B, Sodroski J, Reinherz E. A Soluble CD4 Protein Selectively Inhibits HIV Replication and Syncytium Formation. Nature. 1988;331:78-81.
8. Helseth, E, Kowalski, M, Gabuzda, D, Olshevsky, U, Haseltine WA, Sodroski JG. Rapid Complementation Assays Measuring replicative potential of HIV-1 envelope glycoprotein mutants. J. Virology 1990;64:2416-2420.
9. Olshevsky U, Helseth E, Furman C, Li J, Haseltine WA, Sodroski JG. Identification of Individual HIV-1 gp120 Amino Acids Important for CD4 Receptor Binding. J. Virol. 1990;64:5701-5707.
10. Thali M, Olshevsky U, Furman C, Gabuzda D, Posner M, Sodroski J. Characterization of a Discontinuous HIV-1 gp120 Epitope Recognized by a Broadly Neutralizing Human Monoclonal Antibody. J. Virol. 1991;65:6188-6193.
11. Thali M, Furman C, Wahren B, Posner M, Ho D, Robinson J, Sodroski J. Cooperativity of Neutralizing Antibodies Directed Against the V3 and CD4 Binding Regions of the Human Immunodeficiency Virus gp120 Envelope Glycoprotein. Journal of Acquired Immune Deficiency Syndromes. 1992;6:591-599.
12. Thali M, Furman C, Ho D, Robinson J, Tilley S, Pinter A and Sodroski J. Discontinuous, Conserved Neutralization Epitopes Overlapping the CD4 Binding Region of the HIV-1 gp120 Envelope Glycoprotein. J. Virol. 1992;66:5635-5641.
13. Thali M, Furman C, Helseth E, Repke H and Sodroski J. Lack of Correlation Between Soluble CD4-Induced Shedding of the Human Immunodeficiency Virus Type 1 Exterior Envelope Glycoprotein and Subsequent Membrane Fusion Events. J. Virol. 1992;66:5516-5524.
14. Cao J, Bergeron L, Helseth E, Thali M, Repke H and Sodroski J. Effects of Amino Acid Changes in the Extracellular Domain of the Human Immunodeficiency Virus Type 1 (HIV-1) gp41 Envelope Glycoprotein. J. Virol. 1993;67:2747-2755.
15. Sullivan N, Thali M, Furman C, Ho D and Sodroski J. Effect of Amino Acid Changes in the V1/V2 Region of the HIV-1 gp120 Glycoprotein on Subunit Association, Syncytium Formation, and Recognition by a Neutralizing Antibody. J. Virol. 1993;67:3674-3679.
16. Thali M, Moore JP, Furman C, Charles M, Ho DD, Robinson J and Sodroski J. Characterization of Conserved Human Immunodeficiency Virus Type 1 (HIV-1) gp120 Neutralization Epitopes Exposed Upon gp120-CD4 Binding. J. Virol. 1993;67:3978-3988.
17. Wyatt R, Sullivan N, Thali M, Repke H, Ho DD, Robinson J, Posner M and Sodroski J. Functional and Immunologic Characterization of HIV-1 Envelope Glycoproteins Containing Deletions of the Major Variable Regions. J. Virol. 1993;67:4557-4565.
18. Moore J, Sattentau Q, Wyatt R and Sodroski J. Probing the Structure of the human Immunodeficiency Virus Surface Glycoprotein gp120 with a Panel of Monoclonal Antibodies. J. Virology 1994;68:469-484.
19. Thali M, Charles M, Furman C, Cavacini L, Posner M, Robinson J and Sodroski J. Resistance to Neutralization by Broadly Reactive Antibodies to the HIV-1 gp120 Glycoprotein Conferred by a gp41 Amino Acid Change. J. Virol. 1994;68:674-680.
20. Roben P, Moore J, Thali M, Sodroski J, Barbas C, III and Burton D. Recognition Properties of a Panel of Human Recombinant Fab Fragments to the CD4 Binding Site of gp120 that Show Differing Abilities to Neutralize Human Immunodeficiency Virus Type 1. J. Virol. 1994;68:4821-4828.
21. Ditzel, H, Binley J, Moore J, Sodroski J, Sullivan N, Sawyer L, Hendry RM, Yang W-P, Barbas C, and Burton D. Neutralizing Recombinant Human Antibodies to a Conformational V2- and CD4-Binding Site-Sensitive Epitope of HIV-1 gp120 Isolated Using an Epitope-Masking Procedure. J. Immunol. 1995;154:893-905.

BIOGRAPHICAL SKETCH

Give the following information for the key personnel and consultants and collaborators. Begin with the principal investigator/program director. Photocopy this page for each person.

NAME	POSITION TITLE
Peter E. Wright	Member and Chairman, Dept. of Molecular Biology The Scripps Research Institute

EDUCATION (*Begin with baccalaureate or other initial professional education, such as nursing, and include postdoctoral training.*)

INSTITUTION AND LOCATION	DEGREE	YEAR CONFERRED	FIELD OF STUDY
University of Auckland, New Zealand	B.Sc.	1968	Chemistry
University of Auckland, New Zealand	M.Sc.	1969	Chemistry
University of Auckland, New Zealand	Ph.D.	1972	Chemistry
University of Oxford, England	Postdoctoral	1972-76	Chem/Biochem.

RESEARCH AND PROFESSIONAL EXPERIENCE: Concluding with the present position, list, in chronological order, previous employment, experience, and honors. Key personnel include the principal investigator and any other individuals who participate in the scientific development or execution of the project. Key personnel typically will include all individuals with doctoral or other professional degrees, but in some projects will include individuals at the masters or baccalaureate level provided they contribute in a substantive way to the scientific development or execution of the project. Include present memberships on any Federal Government public advisory committee. List, in chronological order, the title, all authors, and complete reference to all publications during the past three years and to representative earlier publications pertinent to this application. DO NOT EXCEED TWO PAGES.

Professional Experience

New Zealand University Grants Committee Postdoctoral Fellow, Inorganic Chemistry Laboratory, University of Oxford, England, 1972-1973.
Research Associate, Inorganic Chemistry Laboratory, University of Oxford England, 1973-1976.
Lecturer in Inorganic Chemistry, University of Sydney, Australia, 1976-1980.
Senior Lecturer in Inorganic Chemistry, University of Sydney, Australia, 1980-1984.
Member, Department of Molecular Biology, Research Institute of Scripps Clinic, 1984-
Chairman, Department of Molecular Biology, Research Institute of Scripps Clinic, 1987-.

Advisory Positions and Honors

Editor-in-Chief: *Journal of Molecular Biology.*
Editorial Advisory Boards: *Biochemistry, Current Opinion in Structural Biology, Macromolecular Structures, Journal of Biomolecular NMR.*
NSF Biophysics Advisory Panel: 1988.
NIH Molecular and Cellular Biophysics Study Section: 1988; 1989-1993 (Chairman, 1991-1993).
Co-Organizer, Keystone Symposia: "Frontiers of NMR in Molecular Biology," 1991, 1993.
Howard Hughes Medical Institute, Structural Biology Advisory Board, 1993-1996.
Adjunct Professor, Salk Institute, 1992-.
MERIT Award, NIDDK, 1994
Discovery Advisory Board, SmithKline Beecham Pharmaceuticals, 1993-present.

Selected Publications (from a total of 193)

"Low resolution structure of the *Bacillus subtilis* glucose permease IIA domain derived from heteronuclear three-dimensional NMR spectroscopy," W.J. Fairbrother, G.P. Gippert, J. Reizer, M.H. Saier Jr. and P.E. Wright (1992) *FEBS Lett.* **296**, 148-152.
"Folding of peptide fragments comprising the complete sequence of proteins:Models for initiation of protein folding. I. Myohemerythrin." H.J. Dyson, G. Merutka, J.P. Waltho, R.A. Lerner and P.E. Wright (1992) *J. Mol. Biol.* **226**, 795-817.
"Folding of peptide fragments comprising the complete sequence of proteins: Models for initiation of protein folding. II. Plastocyanin." H.J. Dyson, J.R. Sayre, H-C. Shin, G. Merutka, R.A. Lerner and P.E. Wright (1992) *J. Mol. Biol.* **226**, 819-835.
"Relationship between ^{1}H and ^{13}C NMR chemical shifts and the secondary and tertiary structure of a zinc finger peptide." M.S. Lee, A.G. Palmer and P.E. Wright (1992) *J. Biomol. NMR* **2**, 307-322.
"Backbone dynamics of the *Bacillus subtilis* glucose permease IIA domain determined from ^{15}N NMR relaxation measurements," M.J. Stone, W.J. Fairbrother, A.G. Palmer III, J. Reizer, M.H. Saier Jr. and P.E. Wright (1992) *Biochemistry* **31**, 4394-4406.
"Assignment of the aliphatic ^{1}H and ^{13}C resonances of the *Bacillus subtilis* glucose permease IIA domain using double- and triple-resonance heteronuclear three-dimensional NMR spectroscopy." W.J. Fairbrother, A.G. Palmer, III, M. Rance, J. Reizer, M.H. Saier, Jr. and P.E. Wright (1992) *Biochemistry* **31**, 4413-4425.
"Conformation and dynamics of an Fab'-bound peptide by isotope-edited NMR spectroscopy." P. Tsang, M. Rance, T.M. Fieser, J.M. Ostresh, R.A. Houghten, R.A. Lerner and P.E. Wright (1992) *Biochemistry* **31**, 3862-3871.

"Definition of the binding sites of individual zinc fingers in the TFIIIA-5S RNA gene complex." K.R. Clemens, X. Liao, V. Wolf, P.E. Wright and J.M. Gottesfeld (1992) *Proc. Natl. Acad. Sci. USA* **89**, 10822-10826.

"Comparison of backbone and tryptophan sidechain dynamics of reduced and oxidized *E. coli* thioredoxin using ^{15}N NMR relaxation measurements." M.J. Stone, K. Chandrasekhar, A. Holmgren, P.E. Wright and H.J. Dyson (1993) *Biochemistry* **32**, 426-435.

"Characterization of amino acid side chain dynamics in a zinc finger peptide using ^{13}C NMR spectroscopy and time-resolved fluorescence spectroscopy." A.G. Palmer, III, R.A. Hochstrasser, D.P. Millar, M. Rance and P.E. Wright, *J. Amer. Chem. Soc.* **115**, 6333-6345.

"Mapping of the binding interfaces of the proteins of the bacterial phosphotransferase system, HPr and IIA^{glc}." Y. Chen, J. Reizer, M.H. Saier, W.J. Fairbrother and P.E. Wright (1993) *Biochemistry* **32**, 32-37.

"Assignments of ^1H, ^{15}N and ^{13}C resonances, identification of elements of secondary structure and determination of the global fold of the DNA binding domain of GAL4." M. Shirakawa, W.J. Fairbrother, Y. Serikawa, T. Ohkubo, Y. Kyogoku and P.E. Wright (1993) *Biochemistry* **32**, 2144-2153.

"Molecular basis for specific recognition of both RNA and DNA by a zinc finger protein." K.R. Clemens, V. Wolf, S.J. McBryant, X. Liao, P. Zhang, P.E. Wright and J.M. Gottesfeld (1993) *Science*, **260**, 530-533.

"Solution structure of the retinoid X receptor α DNA-binding domain: Identification of a novel helix required for cooperative homodimeric DNA binding." M.S. Lee, S.A. Kliewer, J. Provencal, P.E. Wright and R. M. Evans (1993) *Science*, **260**, 1117-1121.

"The solution structure of the Oct-1 POU-specific domain reveals a striking similarity to the bacteriophage λ DNA-binding domain." N. Assa-Munt, R. Mortishire-Smith, R. Aurora, W. Herr and P. E. Wright (1993) *Cell*, **73**, 193-205.

"Peptide models of protein folding initiation sites. 1. Secondary structure formation by peptides corresponding to the G- and H-helices of myoglobin." J.P. Waltho, V.A. Feher, G. Merutka, H.J. Dyson and P.E. Wright, *Biochemistry* **32**, 6337-6347.

"Peptide models of protein folding initiation sites. 2. The G-H turn region of myoglobin acts as a helix stop signal." H.-C. Shin, G. Merutka, J.P. Waltho, P.E. Wright and H. J. Dyson, *Biochemistry* **32**, 6348-6355.

"Peptide models of protein folding initiation sites. 3. The G-H helical hairpin of myoglobin." H.-C. Shin, G. Merutka, J.P. Waltho, L.L. Tennant, H. J. Dyson and P.E. Wright, *Biochemistry* **32**, 6356-6364.

"Determination of local ligand conformations in slowly tumbling proteins by homonuclear 2D and 3D NMR: Application to heme propionates in leghemoglobin." D. Morikis, R. Brüschweiler and P.E. Wright (1993) *J. Amer. Chem. Soc.* **115**, 6238-6246.

"Electrostatic calculations of side-chain pK_a values in myoglobin and comparison with NMR data for histidines." D. Bashford, D.A. Case, C. Dalvit, L. Tennant and P.E. Wright (1993) *Biochemistry* **32**, 8045-8056.

"Investigation of chorismate binding to a catalytic antibody using transferred nuclear Overhauser effects." A.P. Campbell, P.E. Wright, T.M. Tarasow and D. Hilvert (1993) *Proc. Natl. Acad. Sci. USA* **90**, 8663-8667.

"A molten globule intermediate formed early on the kinetic folding pathway of apomyoglobin" P.A. Jennings and P.E. Wright (1993) *Science* **262**, 892-896.

"Characterization of a folding intermediate of apoplastocyanin trapped by proline isomerization." S. Koide, H.J. Dyson and P.E. Wright (1993) *Biochemistry* **32**, 12299-12310.

"NMR evidence for multiple conformations in a highly helical model peptide." G. Merutka, D. Morikis, R. Brüschweiler and P.E. Wright (1993) *Biochemistry* **32**, 13089-13097.

"^1H resonance assignments and secondary structure of the carbon monoxide complex of soybean leghemoglobin determined by homonuclear 2D- and 3D-NMR spectroscopy." D. Morikis, C.A. Lepre and P.E. Wright (1994) *Eur. J. Biochem.*, in press.

"^1H, ^{15}N and ^{13}C resonance assignments for the first three zinc fingers of transcription factor IIIA." X. Liao, K.R. Clemens, J. Cavanagh, L. Tennant and P.E. Wright (1994) *J. Biomol. NMR* **4**, 433-454.

"^1H and ^{15}N resonance assignments and secondary structure of the carbon monoxide complex of sperm whale myoglobin." Y. Theriault, T.C. Pochapsky, C. Dalvit, M. Chiu, S.G. Sligar and P.E. Wright (1994) *J. Biomol. NMR* **4**, 491-504.

"Solution structure of carbonmonoxy myoglobin determined from NMR distance and chemical shift constraints." K. Ösapay, Y. Theriault, P.E. Wright and D.A. Case (1994) *J. Mol. Biol.* **244**, 183-197.

"Stabilization of a type VI turn in a family of linear peptides in water solution." J. Yao, V.A. Feher, F. Espejo, M.T. Reymond, P.E. Wright and H.J. Dyson (1994) *J. Mol. Biol.* **243**, 736-753.

"Three-dimensional structure of a type VI turn in a linear peptide in water solution: Evidence for stacking of aromatic rings as a major stabilizing factor" J. Yao, H.J. Dyson and P.E. Wright (1994) *J. Mol. Biol.* **243**, 754-766.

"Differential side chain hydration in a linear peptide containing a type VI turn." J. Yao, R. Brüschweiler, H.J. Dyson and P.E. Wright (1994) *J. Am. Chem. Soc.*, **116**, 12051-12052.

OTHER SUPPORT

(Use continuation pages if necessary)

FOLLOW INSTRUCTIONS CAREFULLY. Incomplete, inaccurate, or ambiguous information about OTHER SUPPORT could lead to significant delays in the review and/or possible funding of the application. If there are changes in the information after submission, notify the scientific review administrator of the initial review group before the review; if changes occur after the review, notify the appropriate Institute.

Other support is defined as all funds or resources, whether Federal, non-Federal, or institutional, available to the principal investigator/program director (and other key personnel named in the application) in direct support of their research endeavors through research or training grants, cooperative agreements, contracts, fellowships, gifts, prizes, and other means. However, in the case of prizes and gifts, only those that support the specific project must be reported. Key personnel are defined as all individuals who participate in the scientific development or execution of the project. Key personnel typically will include all individuals with doctoral or other professional degrees, but in some projects will include individuals at the masters or baccalaureate level provided they contribute in a substantive way to the scientific development or execution of the project.

Reporting requirements are: for each of the key personnel, describe (1) all currently ***active*** support and (2) all applications and proposals ***pending*** review or award, whether related to this application or not. If the support is part of a larger project, identify the principal investigator/program director and provide the data for the relevant subproject(s). If an individual has no active or pending support, check "None." Use continuation pages as needed to provide the required information in the ***format*** as shown below. Information may be combined as long as the format remains the same. For example, all key personnel who have no other support may be listed on a single page. **DO NOT SEND** in a separate page for each person listed for whom "None" is checked.

Name Carlos F. Barbas, III Active X Pending _____ None _____

a. Source and identifying no. NIH 1R01AI37470-01 P.I. Carlos F. Barbas, III

Title In vitro evolution of human anti-HIV antibodies

b. Your role on project Principal Investigator % Effort 25%

c. Dates and costs of entire project (For renewals, include only the most recent competitive award. List direct and indirect costs separately.)

12/01/94 - 11/31/97 $383,814 (Direct) $255,722 (Indirect)

d. Dates and costs of current year 12/01/94 - 11/31/96 $132,705 (Direct) $79,946 (Indirect)

e. Specific aims of project Apply CDR walking strategy to human anti-HIV-1 antibodies with CD4-binding site, V3 loop, and gp41 specificities. This approach will be utilized to increase neutralizing potency of these antibodies. The ultimate goal is the construction of a highly potent cocktail of antibodies for therapeutic and prophylactic application in HIV-1 infection.

f. Describe scientific and budgetary overlap No overlap.

g. Describe adjustments you will make if the present application is funded (budget, % effort, aims, etc.)

I will reduce my percent effort on R01AI37470 to 20% to reflect the high priority deserved by the current proposal.

OTHER SUPPORT

(Use continuation pages if necessary)

FOLLOW INSTRUCTIONS CAREFULLY. Incomplete, inaccurate, or ambiguous information about OTHER SUPPORT could lead to significant delays in the review and/or possible funding of the application. If there are changes in the information after submission, notify the scientific review administrator of the initial review group before the review; if changes occur after the review, notify the appropriate Institute.

Other support is defined as all funds or resources, whether Federal, non-Federal, or institutional, available to the principal investigator/program director (and other key personnel named in the application) in direct support of their research endeavors through research or training grants, cooperative agreements, contracts, fellowships, gifts, prizes, and other means. However, in the case of prizes and gifts, only those that support the specific project must be reported. Key personnel are defined as all individuals who participate in the scientific development or execution of the project. Key personnel typically will include all individuals with doctoral or other professional degrees, but in some projects will include individuals at the masters or baccalaureate level provided they contribute in a substantive way to the scientific development or execution of the project.

Reporting requirements are: for each of the key personnel, describe (1) all currently ***active*** support and (2) all applications and proposals ***pending*** review or award, whether related to this application or not. If the support is part of a larger project, identify the principal investigator/program director and provide the data for the relevant subproject(s). If an individual has no active or pending support, check "None." Use continuation pages as needed to provide the required information in the ***format*** as shown below. Information may be combined as long as the format remains the same. For example, all key personnel who have no other support may be listed on a single page. **DO NOT SEND** in a separate page for each person listed for whom "None" is checked.

Name Carlos F. Barbas, III Active _____ Pending X None _____

a. Source and identifying no. NIH Program Project - #PO1 CA27489-16 P.I. Richard A. Lerner, M.D.

Title Consequences of Endogenous Retroviral Expression

b. Your role on project Principal Investigator on one (1) project % Effort 25%

c. Dates and costs of entire project (For renewals, include only the most recent competitive award. List direct and indirect costs separately.)

12/01/94 - 11/31/99 $932,994

d. Dates and costs of current year 12/01/94 - 11/31/95 $170,426

e. Specific aims of project Evolve synthetic antibodies to act as cysteine, serine and metallo-peptideases. Catalyst will be characterized at the amino acid level and kinetic and structural experiments will be performed to determine the mechanism of catalysis.

D scribe scientific and budgetary overlap No overlap.

. Describe adjustments you will make if the present application is funded (budget, % effort, aims, etc.)

No adjustments

OTHER SUPPORT

(Use continuation pages if necessary)

FOLLOW INSTRUCTIONS CAREFULLY. Incomplete, inaccurate, or ambiguous information about OTHER SUPPORT could lead to significant delays in the review and/or possible funding of the application. If there are changes in the information after submission, notify the scientific review administrator of the initial review group before the review; if changes occur after the review, notify the appropriate Institute.

Other support is defined as all funds or resources, whether Federal, non-Federal, or institutional, available to the principal investigator/program director (and other key personnel named in the application) in direct support of their research endeavors through research or training grants, cooperative agreements, contracts, fellowships, gifts, prizes, and other means. However, in the case of prizes and gifts, only those that support the specific project must be reported. Key personnel are defined as all individuals who participate in the scientific development or execution of the project. Key personnel typically will include all individuals with doctoral or other professional degrees, but in some projects will include individuals at the masters or baccalaureate level provided they contribute in a substantive way to the scientific development or execution of the project.

Reporting requirements are: for each of the key personnel, describe (1) all currently ***active*** support and (2) all applications and proposals ***pending*** review or award, whether related to this application or not. If the support is part of a larger project, identify the principal investigator/program director and provide the data for the relevant subproject(s). If an individual has no active or pending support, check "None." Use continuation pages as needed to provide the required information in the ***format*** as shown below. Information may be combined as long as the format remains the same. For example, all key personnel who have no other support may be listed on a single page. **DO NOT SEND** in a separate page for each person listed for whom "None" is checked.

Name Carlos F. Barbas, III Active _____ Pending X None _____

a. Source and identifying no. NIH Program Project IV P.I. Richard A. Lerner, M.D.
Carlos F. Barbas, III

Title In vitro selection of catalytic antibodies

b. Your role on project Principal Investigator % Effort 25%

c. Dates and costs of entire project (For renewals, include only the most recent competitive award. List direct and indirect costs separately.)

05/01/95 - 04/31/00 $932,994

d. Dates and costs of current year 5/1/95 - 4/30/96 $170,246

e. Specific aims of project Develop catalytic antibodies in vitro selection. Develop antibodies which hydrolize peptides.

f. Describe scientific and budgetary overlap None

g. Describe adjustments you will make if the present application is funded (budget, % effort, aims, etc.)

None

RESOURCES AND ENVIRONMENT

FACILITIES: Mark the facilities to be used at each performanc site listed in Item 9, Face Page, and briefly indicate their capacities, pertin nt capabilities, relative proximity, and extent of availability to the project. Use "Other" to describe the facilities at any other performance sites listed in Item 9 on the Face Page and at sites for field studies. Use continuation pages if necessary. Include an explanation of any consortium/contractual arrangements with other organizations.

[X] Laboratory: The proposed research will be performed in a 2,000 sq. ft. laboratory space fully equipped for molecular biology. For molecular biology, we have a designated room for gel electrophoresis and DNA sequencing a variety of heat blocks for eppendorf tubes, a 37°C room equipped with orbital shakers, a Perkin Elmer DNA PCR cycler, a Molecular Devices microplate reader, FPLC apparatus for protein purification, Pharmacia Phast system, Designated area for use of isotopes and software for DNA sequence, reading, recording and manipulating. This instrument is borrowed from another lab. We have UV spectrophotometer. We also have a BIAcore instrument from Pharmacia to characterize binding interactions.

[ ] Clinical:

[X] Animal: The Scripps Research Institute has a managed vivarium with staff available for immunizing and bleeding animals on a fee basis.

[X] Computer: The Scripps Research Institute has a network consisting of over 50 Sun workstation, 8 silicon graphics workstations, 3 VAX minicomputers, 2 Starden graphics supercomputers, a Convex C240 supercomputer and a Cray-Y-MP2E/232 supercomputer. The laboratory is equipped with a Sun workstation and several McIntosh computers.

[X] Office: The P.I. has office space equipped with a McIntosh computer.

[ ] Other (_____):

MAJOR EQUIPMENT: List the most important equipment items already available for this project, noting the location and pertinent capabilities of each.

The Department of Molecular Biology has a Aviv 61DS CD Spectrometer; Applied Biosystems Automated Peptide Synthesizer; DNA Sequencer; 200Hz, 300MHz, 500MHz Bruker NMR Spectrometers; VG ZAB-VSE Mass Spectrometer, Convex C240 computer; Cray and 750 Research XMP-EA 116/SE supercomputer.

ADDITIONAL INFORMATION: Provide any other information describing the environment for the project. Identify support services such as consultant, secretarial, machine shop, and electronics shop, and the extent to which they will be available to the project.

Support services provided by The Scripps Research Institute which are available for the project without additional charge include: Full secretarial services, copy machines, laser printer, fax machines, unlimited telephone service, and Science and Medical library. There are several centers for major equipment which are run by highly qualified support staff who are freely available for consultation at any time. These include: NMR Center, Mass Spectrometry Center, Research Computing Center, and Core Sequencing Service.

SPECIFIC AIMS

This proposal focuses on the development of a new class of therapeutic proteins. *In vitro* evolution techniques will be applied to generate zinc finger proteins which bind specifically and with high affinity to conserved HIV-1 DNA sequences. We will investigate the utility of known cellular uptake and nuclear targeting sequences to deliver extracellular zinc fingers to the nucleus. We will also apply selective techniques to optimize the efficiency of these sequences as well as to attempt to discover new ones. This strategy would allow the protein itself to be used as a drug to selectively modulate transcription and would avoid the use of a gene therapy approach to delivery. The result of this work will be a cocktail of evolved zinc finger proteins which are effective when delivered as soluble proteins or when expressed within a cell and act to block HIV-1 transcription and replication by sequence-specific recognition of conserved regions in the HIV-1 genome. This research will result in the generation of a collection of zinc finger domains which can be assembled to recognize any given stretch of DNA sequence, provide proteins for therapeutic and prophylactic application in HIV-1 infection, and provide zinc finger proteins for NMR investigations of molecular recognition in protein/DNA complexes.

The specific aims are:

1. Continue development of *in vitro* evolution methods for the selective modification of zinc finger affinity and specificity including:

(i) sequential optimization of libraries of zinc finger proteins for the construction of polydactyl proteins which bind with high affinity and specificity to 4 conserved sites in the HIV-1 genome.

(ii) select a set of single zinc finger domains with specificity and affinity for each of 64 possible trinucleotide sites. Demonstrate that this set of preselected zinc finger domains can be utilized to rapidly construct sequence-specific proteins that bind any defined sequence.

2. Construct, characterize, and optimize fusion proteins consisting of evolved zinc finger domains and cellular uptake and nuclear targeting domains for their ability to reach the nucleus of a cell. Demonstrate that these proteins when provided to the culture medium can selectively up or down regulate the transcription of a reporter gene. Study the relationship between protein affinity for DNA, extracellular uptake, and ability to inhibit or enhance transcription.

3. Characterize the efficacy of evolved proteins to inhibit HIV-1 transcription and replication using reporter genes and p24 production assays. Evolved proteins will be:

(i) provided extracellularly as fusions with uptake and targeting sequences.

(ii) intracellularly expressed as envisioned in a gene therapy application.

4. Study the structural basis for sequence-specific recognition of DNA by evolved zinc finger proteins in collaboration with Dr. Peter Wright.

BACKGROUND AND SIGNIFICANCE

Recent studies of HIV-1 infection have characterized it as an incredibly dynamic process involving an unprecedented turnover of both virus and $CD4^+$ lymphocytes(1,2). Plasma virus and virus producing cells were demonstrated to have remarkably short half-lives of approximately two days. These studies indicate that the vast majority of circulating plasma virus is derived from continuous rounds of *de novo* infection, replication, and cell turnover. These studies were made possible by potent new drugs which target the HIV-1 reverse transcriptase and protease for inhibition. These drugs work by blocking new rounds of infection and do little if anything to inhibit the production of virus from previously infected cells. A direct consequence of this mechanism of action is the rapid development of drug resistance.

These studies provide tremendous insight into the disease as well as highlight the limitations of current antiviral drugs. An outcome of these studies is an impetus for new drug discovery efforts which focus on the inhibition of viral replication. Indeed, it has been suggested that immunological reconstitution should be possible even late in the course of the disease if viral replication can be controlled(1).

Inhibition of viral replication has long been recognized as an important goal in HIV-1 therapy. This has been the driving force behind the use of antisense oligonucleotides in HIV-1 therapy(3,4). Other strategies which are well advanced are based on the concept of intracellular immunization(5). These include ribozymes(6,7) and dominant-negative approaches(8). These approaches are most successful when they target essential transcripts which are produced at a low level. Often their function in intracellular immunization strategies is very dependent on achieving high levels of intracellular expression(6,8). This is in contrast to control at the DNA level which presents obvious advantages since only a single site needs to be occupied as compared to targeting the many RNA messages which might be produced from a single gene. While to some degree this may be achieved with oligonucleotides which target homopurine-homopyrimide tracks for triple-helix formation the approach is limited in the sequences which may be targeted(4).

From the simplest of organisms to the most complex, proteins that bind nucleic acids control the expression of genes(9). A distinct advantage of proteins over oligonucleotides is that they can modulate transcription up or down by binding to sites within promoter or coding regions or to sites removed a great distance from these regions(9-13). Practically, activation could stimulate the expression of natural proteins to fight disease while repression could shut down HIV-1 transcription and hence replication. If proteins could be produced and delivered which bind with sequence specificity to stretches of DNA sixteen base pairs or greater in length, they would be able to uniquely target and modulate the expression of a single gene within the human genome. It has now become apparent, based on our understanding and ability to manipulate zinc finger protein domains, that this is an achievable goal.

The zinc finger domain was first observed in the transcription factor IIIA of *Xenopus laevis* (TFIIIA) and has since been identified in over 220 proteins which together contain more than 1300 zinc fingers(14,15). The TFIIIA zinc finger sequence motif, Cys-X_{2-4}-Cys-X_3-Phe-X_5-Leu-

X_2-His-X_{3-5}-His, encodes a modular recognition domain stabilized by the tetrahedral coordination of a zinc ion with the conserved cysteines and histidines. The structure of this motif was first determined using NMR methods in 1989 (16) and shown to consist of an antiparallel b-ribbon packed against an a-helix. Subsequent studies revealed the conservation of this fold in finger motifs of different sequence (15). Naturally occurring proteins generally contain multiple repeats of the zinc finger motif which suggests a modular nature which is unique among the classes of DNA binding proteins. Polydactyl arrays of as many as 37 zinc finger domains allow this modular recognition domain to recognize extended asymmetric sequences(14,15). Practically, modularity of the zinc finger domain has been demonstrated with construction of new three finger complexes(17) Manipulations of this motif have been guided by the X-ray crystal structures of the zinc fingers of the murine transcription factor Zif268(18), the human GLI oncogene (19), and the Drosophila transcription factor Tramtrack (TTK)(20) in complex with DNA. These structures reveal that the individual finger domains contact DNA in the major groove with specific contacts to the DNA from the amino-terminal part of the helix. This class of proteins is distinguished from other classes of DNA binding proteins not only in its modularity but in its ability to recognize RNA as well as DNA(21). Indeed, zinc fingers have even demonstrated sequence-specific recognition of DNA-RNA hybrids(22). Preliminary studies demonstrate that using the phage display approach, we can selectively change both the sequence specificity and affinity of this protein for DNA(23,24). This is also supported by the work of other laboratories(25-27). In our laboratory, preliminary experiments have demonstrated that the modularity of the zinc finger domain allows for as many as six selected domains to be tethered together to create a new protein which recognizes 18 base pairs of DNA sequence. Together these results imply that we will be able to construct proteins which bind with specificity to any site in the human genome.

With respect to the issue of delivery, existing approaches allow for the introduction of genes into stem cells(7). It is reasonable to assume that introduction of a gene encoding the proper zinc finger protein could produce cell populations which do not support viral replication. To some degree this has already been demonstrated. It has recently been reported that transient expression in murine cells of a zinc finger protein engineered to bind a new 9 base pair site could activate a reporter gene or in a construct lacking an activation domain, inhibit the transcription of an oncogenic sequence(28). Another report has simply designed a fusion between two zinc fingers of Zif268 and another transcription to create a protein which binds 10 base pairs(29). They also demonstrated activation of a reporter gene in transient expression experiments.

A strategy which could make this approach more broadly accessible to the HIV-1 infected population and more applicable to other human ailments would be one based simply on administration of the protein. There is a growing body of literature which suggests this should be possible. Three recent examples have demonstrated that the proteins Tat from HIV-1(30), Tax from HTLV-1(31), and human lactoferrin(32) can enter cells when provided extracellularly and find their way to the nucleus to activate transcription. It is therefore reasonable to suspect that engineered proteins could be produced which function as extracellular transcription factors as these natural proteins do. Indeed, conjugates of Tat peptides with proteins such as beta-galactosidase have been demonstrated to be taken up by cells and delivered to the nucleus(33).

As with any approach which attempts specific inhibition in HIV-1, the mutation rate and genetic diversity found *in vivo* present formidable obstacles. Despite this, there are regions within the HIV-1 genome which are conserved over all clades and provide suitable targets to explore this approach(34). The potential to directly and specifically control the expression of genes with proteins, and the application of this strategy to HIV-1 deserves attention.

PRELIMINARY STUDIES

1. Manipulating Proteins In Vitro: Phage display of Antibody Libraries

For several years we have been developing methodologies for manipulating vast libraries of antibodies. Since selective procedures are generally more efficient than screening procedures we sought to mimic the immune systems linkage of recognition and replication, or phenotype and genotype, and to express antibody fragments on the surface of bacteria or phage. This would allow selection of specific antibodies based on their ability to bind to immobilized antigen. To this end we constructed systems for the phage display of Fab fragments(35,36). The most successful of these is the phagemid system pComb3(36).

The display of the antibody (or other proteins) on the surface of the phage allows for the selection of clones by panning against antigen in ELISA wells. This is analogous to an affinity chromatography step. Antigen immobilized on beads or whole cells may also be utilized. After vigorous washing, the bound phage which is now enriched for those bearing antigen-specific Fabs, is eluted with acid or antigen. This phage is then amplified and reselected by further rounds of panning. Each step selects for antigen-specific clones as well as for clones of the highest affinity. In this way one can rapidly generate a panel of antigen-specific Fabs. A single panning step can enrich for specific phage by 10^3 to 10^5 fold. The efficiency of the selection process is due in part to over sampling of the library. This is a distinct advantage of phage since 10^{12} phage in a volume of 50 microliters may be applied to a single ELISA well. Thus for a library of a million clones, each type will be present on a million phage. The method also allows sorting of clones based on affinity as shown by a 250-fold enrichment of a tight binding clone (apparent Ka approx. 10^9 M^{-1}) relative to a weaker binding clone (10^7 M^{-1}) in a mixture of the two following one round of panning. This is a distinct advantage of the monovalent display of the Fab on the phage surface with the pComb 3 system. A monovalent display avoids the avidity effects which accompany the display of multiple copies on the phage surface. This system and others have been utilized in our laboratory and others to clone human antibodies from immune individuals of many different specificities(37). This methodology forms the basis of an annual Cold Spring Harbor Laboratory course.

It is not appropriate here to review on a case-by-case basis how we have utilized this system to produce interesting antibodies. However, I will give several examples which I hope will convince the reviewers that this laboratory has considerable experience in protein engineering and a strong commitment to the development of anti-HIV-1 strategies. In creating novel antibodies without immunization, we were the first to demonstrate that using a synthetic antibody repertoire we could select human antibodies which bind virtually any antigen(see 38

and 37 for a review of the area). We have also utilized this approach to develop catalytic human antibodies(39) with the long-term goal of developing antibodies which could catalytically inactivate viruses by hydrolysis of their envelope proteins. Combining a knowledge of protein structure and function with the phage display approach we developed a methodology which allows for the direct design and selection of human antibodies reactive with human receptors (40,41). With this methodology we have created human Fabs without immunization which bind self antigens, the human integrins $\alpha_{IIb}\beta_3$ and $\alpha_v\beta_3$, with very high affinity, 10^{-10}M. In cloning anti-viral antibodies from humans, we have developed two antibodies which are moving rapidly toward human testing. The first of these is directed against respiratory syncytial virus, an important pediatric pathogen and in immunosuppressed individuals(42). We have recently demonstrated the efficacy of this antibody in animals in collaboration with Bob Chanock of the NIAID(43). The second is IgG1-b12 (the Fab fragment of this antibody is known as HIV-4 or b4/12) which was cloned by the PI and has been characterized in collaboration with Dennis Burton here at Scripps(44,45). This antibody is directed against the CD4-binding site on gp120 and has recently been demonstrated to broadly and potently neutralize primary clinical isolates of HIV-1(46). We have also gone on to optimize the affinity of this promising antibody. In a recent report we improved the affinity 8-fold and demonstrated a correlation of potency in infectivity reduction assays with affinity(47). Further, we demonstrated that the improved antibody neutralized some primary isolates not neutralized by the parental antibody. In the most recent extension of this work we have investigated methodologies for the creation of very high affinity variants of this human antibody. Five libraries of b4/12 were constructed by saturation mutagenesis of complementarity-determining regions (CDRs). Libraries of antibody Fab fragments were displayed on the surface of filamentous phage and selected *in vitro* for binding to immobilized gp120. Sequential and parallel optimization strategies of CDRs were examined. The highest affinity Fab prepared using these strategies was improved 420-fold in affinity. The affinity of this Fab was 15pM as compared to 6.3nM for b4/12. To our knowledge, this is the highest affinity antibody/protein interaction yet described. The methodology we have developed provides a route for the improvement of the affinities of antibodies typical of tertiary immune responses into the picomolar range. It is our hope that these improvements will have profound effects on the utility of this antibody as an HIV-1 therapeutic and prophylactic agent. This work has recently been submitted and is provided in the supplementary material section(48). Together with Dennis Burton we are in the process of establishing, in the Scripps P3 facility, neutralization assays with laboratory adapted and primary viruses. We should have these assays established by the fall of '95. This capability is important since we are currently applying this strategy to a number of antibodies directed against other gp120 and gp41 epitopes with the goal of preparing a cocktail of evolved antibodies for HIV-1 therapy and prophylaxis. Additional references to our on-going anti-HIV antibody project may be found in the biographical sketch of the PI.

2. Changing the specificity and affinity of zinc finger proteins by *in vitro* selection

In 1991, having recognized the power of the phage display approach we had developed for manipulating antibody specificity and having observed the report of the first x-ray structure of

the transcription factor Zif268 in complex with DNA, we proposed in a review article the potential of this technique to modify zinc finger specificity (49). We have pursued this proposal. The results we have obtained make it clear that we will be able to construct proteins which bind any DNA sequence and ultimately produce a family of proteins which act as defined genetic switches which can be utilized to regulate gene expression up or down.

Our studies have involved the *in vitro* evolution of the DNA binding domains of Zif268. Though the three zinc fingers of Zif268, simply referred to as Zif268 in the rest of the text, are of murine origin, they are virtually identical to the human tumor suppressor gene Wilm's tumor protein(50) which in fact binds the same consensus sequence. This is not surprising since the zinc finger domain is highly conserved from yeast to humans and is perhaps the most utilized of all nucleic acid binding motifs in eukaryotes. It has recently been estimated that as much as 1 percent of the DNA in human cells encodes zinc fingers (14). Amino acid sequence analysis of over 1300 finger motifs reveals a well conserved framework with diversity focused in the region demonstrated to contact DNA in the Zif268 structure. This high degree of conservation makes it very difficult if not impossible to generate antibody responses against zinc finger domains and is an important characteristic of this motif which makes it suitable for consideration as a therapeutic protein.

Phage display of the Zif268 protein was achieved by modification of our antibody phagemid display system pComb3. The vector was modified for zinc finger display by excision of the DNA encoding the light chain operon to yield the modified vector pComb3.5. The Zif268 sequence from pzif89 was tailored by PCR for insertion between the Xho I and Spe I sites of pComb3.5(**Fig. 1**). Insertion at these sites results in the fusion of Zif268 with the carboxy terminal segment of the filamentous phage coat protein III. Additional SacI/XbaI sites were included in the insert since these sites allow for the construction of libraries an order of magnitude greater that can be achieved using the XhoI/SpeI sites. In control panning experiments we demonstrated that Zif268 (**Fig. 2A**) and the first three DNA binding fingers of TFIIIA could be functionally displayed on the surface of filamentous phage and maintain specificity for their consensus binding sequences. In these experiments the DNA was immobilized in microtiter wells for phage selection. Immobilization of DNA was facilitated by the design of stable hairpin sequences which present the duplex DNA target of the fingers within a single oligonucleotide which was amino labeled (**Fig. 2B**). The amino linker allowed for covalent coupling of the hairpin sequence to acetylated BSA which was then immobilized for selection experiments by adsorption to polystyrene microtiter wells. Biotinylated hairpin sequences worked equally well for selection following immobilization to a streptavidin coated plate and are now our preferred method of immobilization.

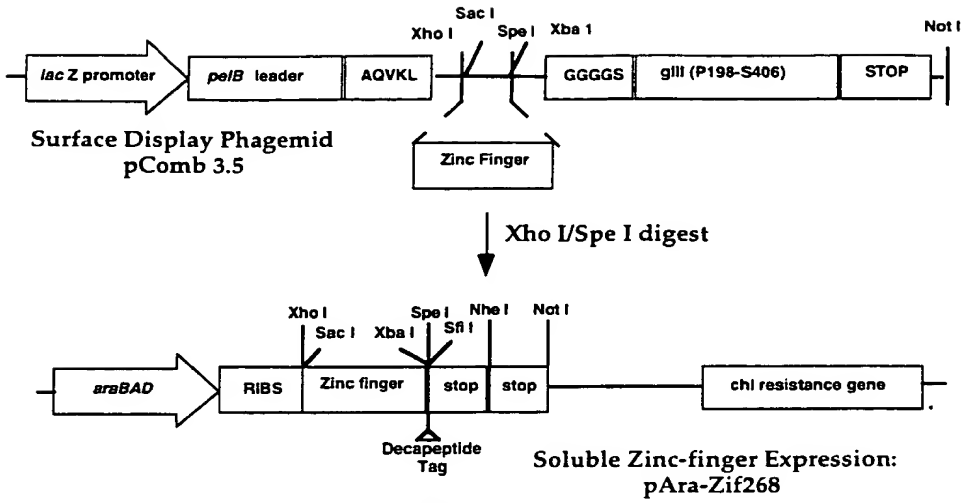


Figure 1.

Libraries of each of the three fingers of Zif268 were independently constructed using the overlap PCR mutagenesis strategy we had applied to antibodies(38). Randomization was limited to six positions due to constraints in the size of libraries which can be routinely constructed. We generally construct libraries with diversities of 10^8 to 10^9 independent clones, though this takes experience (51). Zinc finger protein recognition of DNA involves an antiparallel arrangement of protein in the major groove of DNA, i.e. the amino terminal region is involved in 3' contacts with the target sequence whereas the carboxy terminal region is involved in 5' contacts. Each finger of Zif268 recognizes a nonoverlapping 3 base pair site. It has been speculated based on the subsequently determined structure of the human GLI oncogene (19) that there may be a 1bp overlap region between fingers of zif268. This can be experimentally determined for finger one by modifying the 3' base following its classical 3bp recognition site. We have performed this experiment with all 4 possible base pairs at this position and quantitatively studied the binding kinetics to the 4 duplex DNAs. Binding was absolutely identical indicating the binding sites are nonoverlapping as originally determined. This result is further supported below by work done in our lab and others. Within a given finger/DNA subsite complex, contacts remain antiparallel where in finger 1 of Zif268, guanidinium groups of Arg at helix positions -1 and 6 hydrogen bond with the 3' and 5' guanines, respectively of the GCG target sequence. Contact with the central base in a triplet subsite sequence by the side chain of the helix position 3 residue is observed in finger 2 of Zif268, fingers 4 and 5 of GLI, and fingers 1 and 2 of TTK. Within the three reported crystal structures of zinc-finger/DNA complexes direct base contact has been observed between the side-chains of residues -1 to 6 with the exception of 4(18-20). Based on these observations, residues corresponding to the helix positions -1, 2, 3, 4, 5, and 6 were randomized in the finger 1 and 3 libraries. The Ser of position 1 was conserved in these experiments since it is well conserved at this position in zinc finger sequences in general and completely conserved in Zif268(18). In the finger 2 library, helix positions -2, -1, 1, 2, 3, and 4 were randomized to explore a different mutagenesis strategy where the -2 position is examined since both Zif268 and GLI structures reveal this position to be involved in phosphate contacts and since it will have a context effect on the rest of the domain. Residues 5 and 6 were fixed since the target sequence TTG retained the 5' thymidine of the wild type TGG site. Introduction of ligated

DNA by electroporation resulted in the construction of libraries consisting of 2×10^9, 6×10^8, and 7×10^8 independent transformants for finger libraries 1, 2, and 3, respectively. Each library results in the display of the mutagenized finger in the context of the two remaining fingers of wild-type sequence.

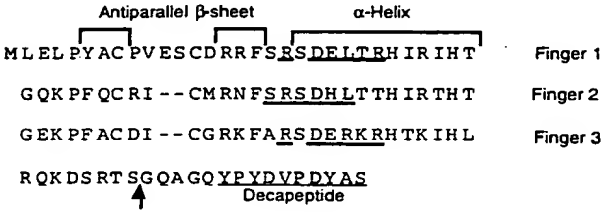


Fig. 2A

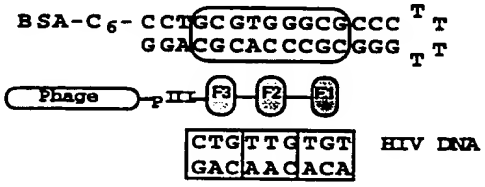


Fig. 2B

In order to examine the potential of modifying zinc-fingers to bind defined targets and to ultimately examine their potential in gene therapy, a conserved sequence within the HIV-1 genome was chosen as our initial target sequence. The 5' leader sequence of the HIV-1 HXB2 clone at positions 104 to 121 relative to the transcriptional initiation start site represents one of several conserved regions within HIV-1 genomes(6,7,34). This is the same region targeted by Flossie Wong-Staal's group for ribozyme cleavage. In these initial experiments we have targeted the 9 base pair region, 113 to 121, shown in **Fig. 2B**.

Following 5 rounds of selection for binding the native consensus or HIV-1 target sequences, functional zinc fingers were rapidly identified with an immunoscreening assay. Expression of the selected proteins in a pAraHA derivative resulted in the fusion of the mutant Zif268 proteins with a peptide tag sequence recognized by a monoclonal antibody (**Fig. 2A**). Binding was determined in an ELISA format using crude cell lysates(**Fig. 3**). A qualitative assessment of specificity can also be achieved with this methodology which is sensitive to at least 4-fold differences in affinity. Several positive clones from each selection were sequenced. Finger 1 selection with the consensus binding site GCG revealed a strong selection for Lys at position -1 and Arg at position 6. Covariation of amino acids of complementary charge between positions -1 and 2 is observed in three clones which contain Lys and Cys at these positions respectively. Clone C7 was preferentially enriched in the selection based on its occurrence in 3 of the 12 clones sequenced. Selection using the HIV-1 target sequence in this region, TGT, revealed a diversity of sequences with a selection for residues with hydrogen-bonding side chains in position -1 and a modest selection for Gln at position 3. Finger 2 selection using the consensus TGG subsite showed a selection for an aromatic residue at -1 whereas selection using the HIV-1 target TTG demonstrated a selection for a basic residue at this position. The preference for Ser at position 3 may be relevant in the recognition of thymidine. Contact of thymine with Ser has been observed in the GLI and TTK structures (19,20). Other modest selections towards consensus residues could also be observed. The NNK doping strategy used to mutagenize the fingers provides all amino acids in 32 possible codons with a single suppressible amber codon. Selections were performed utilizing a supE strain of *E. coli* which resulted in the reading of the amber codon TAG as a Gln during translation. Of the 53 sequences examined, 14 clones possessed a single amber codon. No clones possessed more than one amber codon. Selection

for suppression of the amber stop codon in supE strains has been noted in other DNA binding protein libraries and likely improves the quality of the library since this residue is frequently used as a contact residue in DNA binding proteins (52).

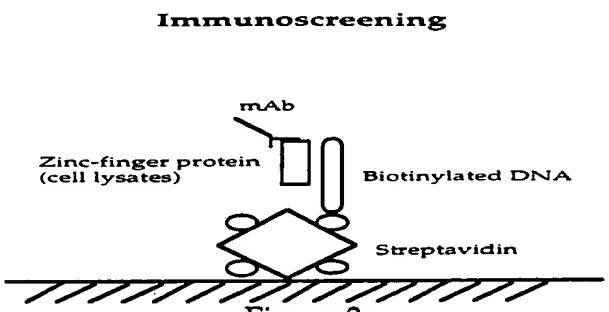


Figure 3

For detailed characterization, high level expression of zinc finger proteins was achieved using the T7 promoter. Purification was facilitated by the isolation of inclusion bodies and refolding of the protein during heparin-sepharose chromatography. Recently, we have moved to purifying the fingers directly from the pAraHA vector using the monoclonal antibody specific for the HA peptide tag. This affords a more rapid route to pure protein without subcloning and refolding steps.

In order to gain insight into the mechanism of altered specificity or affinity we have determined the kinetics of binding using real-time changes in surface plasmon resonance with the BIAcore instrument. The kinetic constants and calculated equilibrium dissociation constants of several of the 11 proteins examined are shown in **Table 1**. The calculated equilibrium dissociation constants for Zif268 binding to its consensus sequence in the form of the designed hairpin or a linear duplex lacking the tetrathymidine loop are virtually identical suggesting that the conformation of the duplex sequence recognized by the protein is not perturbed in conformation within the hairpin. As a measure of specificity, the affinity of each protein was determined for binding to the native consensus sequence and a mutant sequence in which one finger subsite had been changed. Clone C7 is improved 13-fold in affinity for binding the wild-type sequence GCG. The major contribution to this improvement in affinity is a 5-fold slowing of the dissociation rate of the complex . Specificity of the C7 protein is also improved 9-fold with respect to the HIV-1 target sequence. This result suggests that additional or improved contacts are made in the complex and that the binding characteristics of even the natural proteins may be greatly improved. Studies of protein C9 demonstrate a different mechanism of improved specificity. In this case the overall affinity of C9 for the GCG site is equivalent to Zif268 but the specificity is improved 3-fold over Zif268 for binding to the TGT target site by an increase in the off-rate of this complex.
Characterization of proteins modified in the finger 2 domain and selected to bind the TTG subsite reveal the specificity of this finger is amenable to modification. Proteins G4 and G6 bind an oligonucleotide bearing the new subsite with affinities equivalent to Zif268 binding its consensus target. Specificity of these proteins for the target on which they were selected to bind is demonstrated by an approximately 4-fold better affinity for this oligonucleotide as compared to the native binding site which differs by a single base pair. This level of discrimination is similar to that reported for the native finger 1 and mutants (26). The finger 3

modified protein A14 was selected to bind the native finger 3 subsite and binds this site with an affinity which is only 2-fold lower than Zif268. A14 differs radically in sequence from the native protein in the recognition subsite. Sequence specificity in 10 of the 11 proteins characterized was provided by differences in the stability of the complex. Only a single protein, G6, achieved specificity by a dramatic change in on-rate., Characterization of proteins F8 and F15 demonstrate that the 3 base pair recognition subsite of finger 1 can be completely changed from GCG to TGT and that new fingers can be selected to bind this site. Each of the 11 proteins studied in detail demonstrated a higher affinity for the target on which it was selected. BIAcore studies revealed that changes in the specificity of the fingers are governed in almost all cases by changes in the stability of the protein/DNA complex as reflected in the k_{off} values. Furthermore, different mechanisms for the improvement or modification of Zif268 proteins for the target were observed. Based on off-rate measurements, the $t1/2$ of the Zif268-DNA complex is improved from 58 minutes for the natural protein to 289 minutes for the C7 protein. Measurements of half lives of DNA/protein complexes using electrophoretic mobility shift assays are difficult and have led to some disagreement in the literature(23). We have found that BIAcore analysis is a valuable and extremely reproducible technique for the kinetic examination of protein/DNA interactions.

Zinc finger Protein	Binding site	k_{on} (x 10^4) (M^{-1} s^{-1})	k_{off} (x 10^{-4}) (s^{-1})	K_d (x 10^{-9}) (M)	K_d/K_d(target)
Finger 1					
WT	**GCG**	3.1 ± 0.04	2.0 ± 0.1	6.5	1
	TGT	1.1 ± 0.2	9.0 ± 1.0	81.8	12.6
C7	**GCG**	8.0 ± 0.7	0.4 ± 0.1	0.5	1
	TGT	0.9 ± 0.1	4.9 ± 2.0	54.4	108.8
C9	**GCG**	2.0 ± 0.2	1.3 ± 0.3	6.5	1
	TGT	0.9 ± 0.1	23.0 ± 3.0	255.6	39.3
F8	**TGT**	3.7 ± 1.0	11.0 ± 1.5	29.7	1
	GCG	4.8 ± 0.1	52.0 ± 0.9	108.3	3.6
F15	**TGT**	1.9 ± 0.1	7.9 ± 1.0	41.6	1
	GCG	0.9 ± 0.03	17.0 ± 1.7	188.9	4.5
Finger 2					
WT	**TGG**	3.1 ± 0.04	2.0 ± 0.1	6.5	1
	TTG	1.2 ± 0.2	9.9 ± 0.02	83.7	12.9
G4	**TTG**	3.3 ± 0.2	2.1 ± 0.1	6.4	1
	TGG	2.5 ± 0.6	5.7 ± 0.2	22.8	3.6
G6	**TTG**	10.0 ± 1.0	4.6 ± 0.3	4.6	1
	TGG	0.7 ± 0.1	1.4 ± 0.1	20.0	4.3
Finger 3					
WT	**GCG**	3.1 ± 0.04	2.0 ± 0.1	6.5	1
	CTG	0.9 ± 0.1	8.8 ± 0.03	101.0	15.5
A14	**GCG**	1.3 ± 0.1	1.7 ± 0.0	13.1	1
	CTG	0.2 ± 0.0	10.0 ± 0.4	500.0	38.2

Table 1.

These results demonstrated that each of the three zinc fingers of Zif268 is amenable to modification using an *in vitro* selection strategy based on phage display. Reports published during the course of this work have demonstrated some success in modifying the specificity of a single finger of Zif268 by phage display by targeting the four core recognition residues at positions helix positions -1, 2, 3, and 6 (25,26). They, however, were unable to change the entire three base pair subsite. In a more recent report Klug's group has targeted 7 positions and demonstrated the ability to change the three base pair subsite but not improve upon the binding to the natural finger(27). However in this study, DNA specificity was determined for zinc finger proteins on phage and soluble protein was not studied. We believe this to be an important caveat of their work since we have found it difficult to characterize antibodies on phage reliably and hence only characterize soluble proteins. With our mutagenesis strategy the entire 3 base pair recognition site of a zinc finger has been modified and we have demonstrated the ability to improve the affinity of the natural finger. It is important to note that our success is partly due to the size of the library we were able to construct in combination with a more extensive mutagenesis strategy. The libraries generated in our preliminary studies were up to three orders of magnitude larger than those reported by the other groups.

The crystal structure of the Zif268/DNA complex revealed a simple coding relationship between protein and DNA target(18). To explore the issue of a coding relationship, selections were performed using the native consensus sites as well as new target sites. Our results have implications for the development of a coding relationship between protein sequence and DNA subsite sequence. Selection of the finger 1 library for binding to the wild type binding site GCG produced a strong selection for Arg at position 6. This result is in agreement with the wild type protein and the structure of the protein/DNA complex which reveals direct hydrogen bonding contacts with the 3' guanine of the target site. Position -1 following selection produced a strong selection for Lys even though the wild type protein has an Arg at this position which contacts a 5' guanine via two hydrogen bonds. In the structure of the Zif268 complex, the Arg at position -1 is stabilized by an intramolecular hydrogen bond to Asp at position 2. In our experiment, covariation of residues at positions -1 and 2 was observed in the finger 1 GCG selection. Predominately, residues of complementary charge were selected at these positions as suggested by the structure of the complex. However, the protein of the highest affinity and specificity, C7, did not follow this trend. Characterization of protein A14 derived from the finger 3 GCG selection demonstrated that a finger sequence radically different from the wild type can bind the wild type target with only a 2-fold reduction in affinity. Collectively these data suggest that a single general coding relationship between protein and target DNA sequence does not exist. Multiple coding relationships may exist however which define subsets of finger/ DNA sequence. These results indicate that rational design of efficient new finger proteins based on the small subset of well characterized fingers will not be possible and that selective approaches outlined above will be required. This preliminary work has now been published and is given in reference 23 in the appendix. The details of the BIAcore based characterization of DNA/protein interactions are the subject of a special journal issue (24) now in press devoted to the utility of the technique and is given in the appendix.

These results highlight our ability to modify the specificity of single finger domains. The goal of this work is, however, to target a unique site in the $3.5x10^9$ bp genome of a human. Given the 3bp site recognition of a single zinc finger domain of the Zif268 type, a five finger complex would recognize a single 15bp sequence per $1x10^9$ bp (4^{15}) and a 6 finger complex would recognize a single 18bp site per $6.8x10^{10}$ bp (4^{18}) or 16 times the size of the human genome. These calculations were made with the assumption of random DNA sequence. We side in favor of recognition of a longer sequences to avoid side-effects since minimal lengths will be subject to empirical determination.

Zinc-finger protein	Binding site	k_{on} (x 10^4) ($M^{-1}s^{-1}$)	k_{off} (x 10^{-4}) (s^{-1})	K_d (x 10^{-9}) (M)	K_d/K_d(target)
F12#4	**GCGTTGTGT**	2.5 ± 0.4	0.9 ± 0.0	3.6	1
	GCGTGGTGT	0.5 ± 0.1	4.3 ± 0.0	86.0	24
	GCGTTGGCG	0.8 ± 0.1	5.5 ± 0.0	68.8	19.1
	GCGTGGGCG			NB	
f15F2#3	**GCGTTGTGT**	1.4 ± 0.2	2.8 ± 0.0	20.0	1
	GCGTGGTGT	1.1 ± 0.0	12.1 ± 0.2	110.0	5.5
	GCGTTGGCG	1.0 ± 0.0	12.0 ± 0.2	120.0	6.0
	GCGTGGGCG	0.6 ± 0.1	15.1 ± 0.0	251.7	12.6
f15F2#9	**GCGTTGTGT**	2.9 ± 0.3	0.8 ± 0.0	2.8	1
	GCGTGGTGT	2.3 ± 0.2	5.5 ± 0.1	23.9	8.5
	GCGTTGGCG	1.4 ± 0.3	6.7 ± 0.1	47.9	17.1
	GCGTGGGCG	0.6 ± 0.1	5.6 ± 0.3	93.3	33.3

Table 2.

There are two strategies evident for the construction of new polydactyl complexes, sequential or parallel. Sequential selection results in the conversion of Zif268 one finger at a time to a new 9bp specific protein. This strategy would take into account any changes in DNA structure relative to the Zif268 recognition sequence which may accompany sequence change as well as any finger/finger dependence. Parallel optimization assumes that both DNA subsite and zinc finger domains are modular in function and structure. This approach would allow the rapid construction of polydactyl complexes simply by tethering together preselected fingers. We have now tested both of these approaches by producing novel polydactyl proteins which bind with high affinity and specificity to 6 bp of the HIV target described above. The result are given in **Table 2.** Protein F12#4 was produced by combining the collections of fingers 1 and 2 preselected to bind their respective 3bp DNA subsites followed by reselection to bind an oligo containing 6bp of the HIV DNA sequence, finger 3 is the natural Zif268 finger. Note that this protein binds the desired target with an affinity which is better that observed for Zif268 and in this assay we did not detect binding to the original Zif268 target sequence (different by 4bp). Proteins f15F2#3 and f15F2#9 were constructed by using the finger 1 mutant f15 which binds the HIV subsite TGT which we characterized above and randomizing finger 2 to create another library of 10^9 clones. This library was then selected to bind the target containing the 6bp site. In this case we have now characterized two clones following ELISA screening. Clones bound

to the target with a 10-fold difference in affinity and a modest difference in specificity is noted between the clones. Selection of finger 3 is in progress.

As noted above, we anticipate using proteins containing 5 or 6 fingers. To date, though proteins containing as many as 37 fingers have been cloned, structural data is only available in which 2 or 3 fingers specifically contact DNA. Binding sites have only been determined for a few natural finger containing proteins, none of which recognize the extended length of sequence we intend to target. To further investigate the parallel strategy for the production of many fingered complexes, we have recently constructed 5 proteins wherein we polymerize the C7 finger described above to produce 2,3,4,5, and 6 finger proteins (**Fig. 4**). Individual domains were constructed using the TGEKP linker sequence which is the consensus sequence for the linker region. Note in **Fig 2A** that TGQKP and TGEKP linker sequences are used to connect finger 1 and 2 or 2 and 3, respectively in Zif268. Binding studies of these proteins indicate they specifically bind their 6,9,12,15, and 18bp targets: $(GCG)_{2\text{-}6}$. We have not detected any increase in nonspecific binding in the 6 finger protein 6C7. In fact we observe an increase in specificity with increase in length when we compare binding to the target and a TFIIIA binding sequence as a control. The most interesting result of this study was that half-site recognition was not sufficient to result in binding. For example, while the 3C7 proteins binds the 6C7 target since it has two copies of the 3C7 binding site, the 6C7 protein shows no binding to the 3C7 target which contains half of the 6C7 binding site. This demonstrates that polydactyl proteins recognize with great selectivity their target over targets containing partial binding sites. This is extremely important since the goal of this project is to recognize extended sequences uniquely and at high affinity. This selectivity is likely derived from the energetic cost of having three fingers involved in disruptive interactions with the DNA which prevents or overwhelms formation of the half-site complex. A productive six-finger complex involves continuous contact in the major groove over two helical turns. We have recently purified and analyzed the binding kinetics of the 3 finger protein 3C7. 3C7 binds its target GCGGCGGCG with a 5.5nM affinity as compared to 181nM for the Zif268 target GCGTGGGCG which differs by 2bp. A report by Desjarlais and Berg has described a designed 3 finger protein which recognizes the same sequence as 3C7 but demonstrated an affinity which is 2000-fold poorer than 3C7, 11uM(17). Their protein utilized the same linker sequence but with a designed Cys-X_2-Cys type finger as compared to our selected Cys-X_4-Cys type finger C7. This comparison demonstrates the power of combining selection with protein design and highlights current short-comings in direct design. These results demonstrate that we can construct proteins which recognize at least 18bp of DNA sequence in a specific fashion.

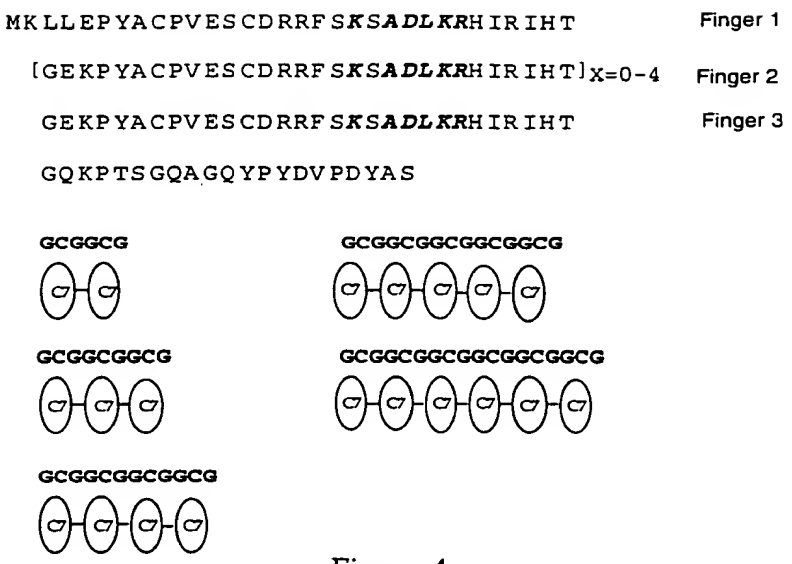


Figure 4

In order to assess the functional properties of the new zinc fingers generated, we have constructed an *E. coli* based *in vivo* system which can serve as a functional screen. This system utilizes two plasmids with different antibiotic resistance genes and compatible replicons colE1 and p15. Cytoplasmic expression of the zinc finger is provided by the arabinose promoter in the colE1 plasmid pAraHA, described above. The p15 replicon containing plasmid contains a zinc finger binding site in place of the lac repressor binding site in a plasmid which expresses the alpha fragment of beta-galactosidase. The binding of the zinc finger to this site on the second plasmid shuts-off the production of beta-galactosidase and thus novel zinc fingers can be assessed in this *in vivo* assay for function using a convenient blue/white selection. In control experiments with Zif268, growth in the presence of arabinose induces expression of the zinc finger. The protein product binds to the zinc finger binding site on the second plasmid and represses the lac promoter. This effectively shuts down beta-galactosidase production and produces only white colonies. No effect is observed with a plasmid with no insert or one expressing the fingers of TFIIIA. This experiment provides further evidence of the potential of this class of proteins to selectively turn-off a gene. This system will allow for the specificity and efficacy of the evolved fingers to be characterized in the context of the 4×10^6 bp genome of E. coli. This is a relevant screen since the concentration of DNA found in E. coli closely approximates the concentration found in the eukaryotic nucleus(9).

RESEARCH DESIGN AND METHODS

Specific Aim 1: Continue development of *in vitro* evolution methods for the selective modification of zinc finger affinity and specificity including:

(i) sequential optimization of libraries of zinc finger proteins for the construction of polydactyl proteins which bind with high affinity and specificity to 4 conserved sites in the HIV-1 genome.

We have selected 4 target sequences in the HIV-1 genome. The criteria for selection were that 1) they be located proximal to the 5' end since early termination will inhibit synthesis of more viral genes 2) the targeted sequence should be unique to HIV-1 and demonstrate a high degree of genetic stability 3) the target sequence should, if possible, be located in an open chromatin configuration to allow the best access to evolved fingers. The third criteria is important since the association of DNA with nucleosomes can confer selective access to DNA. With respect to chromatin studies we are guided by two reports(53,54). The target sequences are given in **Fig. 5**. Consistent with the Los Alamos data base, uppercase letters indicate 100% conservation of nucleotide bases and lower case letters represent bases conserved in at least 50% of sequences. Multidata base searches reveal these to be HIV-1 specific targets. The individual sites are discussed in detail below and proteins will be constructed in the order given in **Fig. 5**.

```
TGT-GCC-CgT-CTG-TTg-TGT   559-576 in HXB2 genome (Wong-Staal seq.)
AAA-TCT-CTA-GCA-GTG-GCG   623-640 in HXB2 genome (PBS region)
AGT-GGC-gag-CCC-TCA-GAT   403-420 in HXB2 genome (Sp1-TATA seq.)
TCC-tAT-GGC-AGG-AAG-AAG   5965-5982 in HXB2 genome (Tat/Rev)
```

Figure 5

As discussed above we have made progress toward construction of a polydactyl finger protein which recognizes bases +104 to +121 in the 5'-nontranslated region; 559-576 in HXB2. This region is located within the DNase I hypersensitive site 3 region defined by positions -65 to +130 and thus should be an accessible site(54). This specific target has been the major focus of Flossie Wong-Staal's ribozyme efforts and has led to promising results(6,7). The two positions shown in lower case indicate diversity in that position. Position +111(566 in HXB2) is A in the clade B isolate CAM1 and consensus O sequences. Otherwise, it is a conserved as G. Position +118(573 in HXB2) is G in clade B with the exception of the MN isolate where it is A. This position is A or C in other clades. To account for this variability we will select fingers which are insensitive to substitution at these positions by simply providing a mixed base at these positions during selection. We are currently performing this experiment with the library described above wherein we have identified and fixed the finger recognizing positions +119 to +121. Indeed natural fingers which are insensitive to substitutions at a given site in the triplet are known (15).

Primer binding site region target. Initiation of HIV-1 reverse transcription occurs by the extension of the Lys3 tRNA primer bound near the 5' end of the genomic RNA designated the primer binding site or PBS (55). The PBS is complementary to the 18 nucleotides of the 3'CCA end of the Lys3 tRNA. The PBS is situated in the 5'-nontranslated region and is part of a stretch of 29 contiguous bases which are conserved over all clades;AAA-ATC-TCT-AGC-AGT-GGC-GCC-CGA-ACA-GGg-AC(primer binding site underlined). Deletion analysis studies indicate retention of the 5' region of the PBS is most essential(56). Deletion of the first 9bp of the PBS does not allow for the production of virus. Deletions which are more 3', slow the kinetics of viral production for a single infection cycle after which revertants containing a complete PBS site and normal kinetics are recovered. It has been suggested that templating with Lys3 tRNA allows for continuous regeneration of 3' mutants as compared to 5' mutants which perhaps do not allow for extension of the tRNA primed template(56). Another study generated a 2bp insertion mutation into the 5' region of the PBS (57). This mutation drastically

decreased the number of viral particles produced. Those particle which were produced were noninfectious. These studies together with the conservation of the region 5' to the PBS may indicate this region has multiple functions in HIV replication. We propose to target the 5' most bases of the PBS and the conserved 13 bps upstream as shown in **Fig. 5** ; +170 to +187. Further this site provides for an 8bp overlap with an AP-3-like site which has recently been defined between positions +162 to +177 and is a potential regulatory element(54). This region is believed to be more accessible to binding proteins following activation of HIV-1 gene expression. Targeting this site, in addition to blocking pol II, may displace an important transcription factor resulting in greater reduction in transcription. Additional highly conserved sites of an appropriate size exist in the DNase I hypersensitive site 4 region (+202 to +266) which could also be important targets. These will be explored given time.

Sp1-TATA region target. The LTR is rich in binding sites for host cell factors(58). Targeting in this region is difficult if one wishes to avoid the common promoter/enhancer elements. Targeting these common regions could have broadly detrimental effects. The Sp1/TATA region falls within the third hypersensitive site defined by positions -64 to -5(53). This is a major constitutive DNase I hypersensitive site and is thus accessible. We have chosen to target the region consisting of the 3' region of the TATA proximal Sp1 site extending toward the TATA binding site given in **Fig. 5**. The region is well conserved across clades with the exception of a single finger site. At this site we will utilize a nonspecific recognition finger which we have previously isolated. There is one sequence reported which shows insertional mutation in this region which will not be covered by this finger, isolate HAN. The significance of this isolate is not known. Occupation of the target site is expected to exclude binding to the TATA proximal Sp1 site and should have a dramatic effect on transcription since Sp1 and Tat are thought to synergistically activate the HIV-1 promoter(59). Linker scanning mutagenesis studies have demonstrated that disruption of the -39 to -22 region, which our target site overlaps, severely cripples LTR driven transcription(60). In other studies, modification of sequence around the TATA box has been shown to diminish activation by Tat(review). This sequence is duplicated in the 3' LTR region within Nef.

Tat/Rev target. As an exception to the 5'-proximal selection criteria we propose a sequence which codes an overlapping region of Tat and Rev given in **Fig. 5**. Again this is a highly conserved region with diversity known for a single base pair which is generally A but has been observed as T or C in consensus O. Targeting this site recognizes the essential role played by these proteins in transcription and replication(55). This region has been successfully targeted in anti-sense strategies(3). This target site is not known to reside in a DNase I hypersensitive region. The extent to which chromatin structure will inhibit access to this site is not known. Further, site of integration and cell type will affect accessibility of the targets described above which are known to reside in accessible regions in the three chronically infected cell lines studied. In a preliminary study by Klug, targeting a coding region in the BCR-ABL fusion oncogene with a low affinity zinc finger resulted in repression of transcription (28). The accessibility of the site had not been previously examined. If targeting of evolved zinc finger proteins is not limited to hypersensitive sites, the potential range of target sequences will be greatly expanded. This point needs to be studied.

Construction of evolved 6-finger proteins will initially follow the strategy outlined in **Fig. 6.**

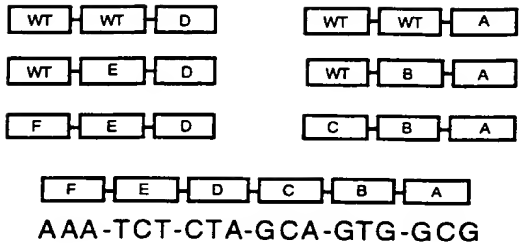


Figure 6

In this approach the 18bp target is divided into two 9bp subsites, ABC and DEF. The sequence shown is the PBS region target. In the first step, fingers A and D are selected in the context of the two remaining natural Zif268 fingers (wt) using the finger 1 libraries which have been constructed, preliminary results and below. Note for site A, we have previously selected an optimal sequence to bind the GCG site, the C7 finger described above. The C7 finger can be used directly so selection in this case begins with A fixed, B randomized, and the final wt finger. Experimentally, for the targets described above, the 9bp subsites are evolved in parallel, involving a total of 5 libraries: A and D which are already constructed, B, C, E, and F. Each library will consist of three sublibraries constructed using different doping oligonucleotides shown in **Fig. 8**. Based on our experience, each library will be constructed by electrotransformation of 18 ligation reactions, 6 per doping strategy used. This will generate between $1\text{-}5\times10^9$ independent clones. In the final step the 9bp subsites are connected using the consensus linker described in the preliminary results section in the construction of the 6C7 protein. Optimization of this linker region could also be performed to further enhance affinity. Phage selections will be performed using biotinylated hairpin oligos as discussed above. Phage will be incubated in solution with the biotinylated target sequence, a nonspecific competitor salmon sperm DNA, and a specific competitor mix. Following equilibration, phage which bind best to the biotinylated oligo are captured with streptavidin. Nonspecific phage are then washed away and specific phage eluted and allowed to infect E. coli as described above. The specific competitor will consist of a randomized target region in the context of the other selected positions. For example, selection of the B subsite requires biotinylated wt-B-A oligo and the specific competitor wt-NNN-A which is non-biotinylated. N represents a mix of the 4 possible bases to give a specific competitor mix which consists of 64 different oligos. The target sequence will be represented at 1/64 in this competitor mix but since we add the specific competitor mix after 2 panning rounds with nonspecific competitor only, we are not in danger of losing a rare clone since the library at this point consists of fewer than 10^5 clones which are represented by 10^{12} phage in each step. Stringent selections are performed to ensure specificity and to minimize the level of nonspecific binding. This maximizes K_N/K_{op} to create an efficient DNA binding protein as discussed in aim 2.

As each new finger pool is selected a subset of clones which recognize the desired target sequence are obtained following the preliminary ELISA screen for specificity. Following sequencing to identify unique clones, several proteins are purified and studied in detail using the BIAcore instrument as described above. The best clone is taken forward for randomization and selection of the next finger. The 6-finger protein will be characterized using the E. coli

blue/white screen, see preliminary results, with appropriate controls (plasmids with target sites for the other 6-finger proteins). Affinity will be characterized with the BIAcore if its dissociation constant is greater than 10^{-12} M or nitrocellulose filter binding assays(61) if it is less. To define in detail all possible binding sites for the 6-finger protein we will select its preferred binding sequences from a pool of random sequence using the selection amplification and binding assay(62). This will tell us what substitutions in the target sequences will be tolerated by the evolved proteins. Based on our experience in optimizing the affinity of antibodies, we are confident that we can do the same with the 6-finger proteins if required. However, considering that half-site affinities will be in the nanomolar range, this will not likely be necessary.

(ii) select a set of single zinc finger domains with specificity and affinity for each of the 64 possible trinucleotide sites. Demonstrate that this set of preselected zinc finger domains can be utilized to rapidly construct sequence specific proteins that bind any defined sequence.
As discussed above in the Preliminary Results section we have made substantial progress in demonstrating that we can manipulate the specificity, affinity, and connectivity of zinc finger domains. The goal of this aim is to prepare a collection of predefined and characterized zinc finger domains from which polydactyl proteins can be rapidly constructed which bind any defined stretch of DNA. This will make this approach easily accessible to the scientific community since the need to construct libraries and select new fingers will not be required. Defined fingers will simply be tethered together from a predefined collection. This will allow us to rapidly construct 6-finger proteins which recognize sequences beyond the targets described so they can be rapidly evaluated for anti-HIV activity. This is feasible since we have demonstrated that zinc finger domains possess a high degree of modularity which allows for the tethering of individual domains to build new proteins which bind up to 18bp of DNA sequences(preliminary results). Since fingers of the Zif268 type bind 3bp nonoverlapping sites, all possible DNA trinucleotide combinations can be defined by the 64 3bp sites shown below. All 64 biotinylated hairpin oligonucleotides will be synthesized and utilized to select from a pool of ~5 x 10^9 mutants of the finger 1 of Zif268.

AAA	ACA	AGA	ATA	TAA	TCA	TGA	TTA
AAC	ACC	AGC	ATC	TAC	TCC	TGC	TTC
AAG	ACG	AGG	ATG	TAG	TCG	TGG	TTG
AAT	ACT	AGT	ATT	TAT	TCT	TGT	TTT
CAA	CCA	CGA	CTA	GAA	GCA	GGA	GTA
CAC	CCC	CGC	CTC	GAC	GCC	GGC	GTC
CAG	CCG	CGG	CTG	GAG	GCG	GGG	GTG
CAT	CCT	CGT	CTT	GAT	GCT	GGT	GTT

Figure 7

-3	-2	-1	1	2	3	4	5	6	7
F	S	R	S	D	E	L	T	R	H
F	S	(NNK)	S	(NNK)	(NNK)	(NNK)	(NNK)	(NNK)	H
F	S	(NNK)	(NNK)	(NNK)	(NNK)	L	(NNK)	(NNK)	H
F	(VNS)	(VNS)	(VNS)	(VNS)	(VNS)	L	(VNS)	(VNS)	H

Figure 8

Finger 1 libraries have been constructed as shown in **Fig. 8**. We have targeted the region between the conserved Phe at position -3 relative to the helix start site and the zinc coordinating His at position 7. This covers the entire range of positions shown to contact DNA and provides for the optimization of flanking context regions. The positions targeted for randomization are shown under the native finger 1 sequence. The first library is as we previously described(23). The next library introduces diversity at position 1 since it will have a context effect as described above (preliminary results) on the residues flanking it. To compensate for increased diversity and maintain a library which can be constructed to be as complete as possible we fix the relatively highly conserved leucine at position 4 which does not contact DNA but is involved in stabilizing the fold of the finger. Note that the NNK dope provides all 20 aa's in 32 codons. A library size of ~10^9 should be prepared to ensure representation of all possible combinations (32^6). The last library is an adaptation and refinement of the mutagenesis strategy of Klug who utilized VNN. V represents a mix of A, G, and C with T excluded. This codes for 16 aa's in 48 codons and eliminates coding of Phe, Trp, Tyr, Cys, and all stops. The excluded residues are less frequently observed at these positions in natural fingers. A VNS dope provides the same 16 aa's encoded by only 24 codons. This decreases the size of the library required for complete representation of all AA combinations by ~128-fold; compare 48^7 to 24^7. This allows us to explore randomization of an additional context position. With this collection of libraries we are confident we can select the requisite pool of 64 fingers. This is supported by our recent selection of fingers which bind TAT and TTT sites which other groups have failed to do and our improvement of the affinity of the natural finger 1 site in the C7 protein which was attempted but not achieved in other reports(26,27). This success is likely do to the fact we have constructed large libraries and have used more extensive mutagenesis. The phagemid system we use facilitates this. The library reported by Klug for example consisted of only 2.6 x 10^6 clones(27). Competitive selections will be performed as described above but using a wt-wt-NNN oligo as a specific competitor. Following selection fingers will be quickly screened in the ELISA format for binding to the 64 oligos in the microtiter format. This will allow us to quickly isolate clones with the highest affinity and selectivity for the 3bp site they were selected to bind. Clones which survive this screen will be sequenced, unique proteins purified, and subjected to BIAcore analysis to further define their binding characteristics and mechanism of base pair discrimination. Fingers which tolerate substitution at a given position can also be identified in this screen and may be important as described above. We will first select fingers to bind the 3bp subsites which are found in our HIV-1 targets and are underlined in **Fig. 7**. Using this collection of defined fingers we will construct, using the strategy described for the 6C7 protein, the two HIV-1 targets which will be constructed first using the sequential selection. The properties of the proteins will then be compared by characterization as described for the proteins derived by sequential selection to HIV-1 sequences. In this manner we will determine the optimal strategy for the construction of polydactyl proteins. Note that even if sequential selection provides some benefits with respect to affinity or specificity, the predefined finger 1 proteins could be utilized directly in the sequential selections since sequential selection of a 6-finger protein defines 6bp using the finger 1 libraries(A and D). We will then use the optimal strategy for the construction of the remaining two 6-finger proteins.

Specific Aim 2: Construct, characterize, and optimize fusion proteins consisting of evolved zinc finger proteins and cell entry and nuclear targeting domains for their ability to reach the nucleus of a cell. Demonstrate that these proteins when provided to the culture medium can selectively up or down regulate the transcription of a reporter gene.

The focus of this aim is the development of evolved zinc finger proteins into a new class of highly specific pharmaceuticals. If we can successfully make the protein itself into a drug, we will be able to bypass any concerns and complications of a gene therapy approach to protein delivery. By tagging our evolved proteins with protein fragments or domains responsible for directing entry into cells and targeting to the nucleus, the proteins themselves would become specific drugs. Selective repression by simple administration of the protein would stop HIV transcription and replication. Alternatively, activation would prompt cells to boost their own levels of protective proteins, proteins which perhaps could not be provided directly as drugs due to stability or site of action requirements, for example an integral membrane protein. Cytokine levels of a specific cell populations could be modulated by engineering cell-type specificity, perhaps by fusion with known ligands, into evolved finger proteins.

To explore the potential of the evolved proteins as drugs without the need to deliver the genes encoding them, fusions with entry and targeting sequences will be studied. This study is based on three recent examples within the literature which demonstrate that Tat from HIV-1 (30), Tax from HTLV-1 (31), and human lactoferrin (32) can enter cells when provided extracellularly and find their way to the nucleus to activate transcription. Indeed in 1991, Mann and Frankel suggested that Tat may be useful for the delivery of proteins into cells(63).

It is appropriate at this point to discuss the issue of dosage of protein required for biological activity. Ptashne has elegantly described the parameters which influence the efficiency of a DNA binding protein(9). In simplistic form the efficacy of a DNA binding protein will be governed by the equation:

$$(O)/(O_t) = K_{op}(N_t)/K_N(R_t)$$

where (O) and (O_t) = concentration of free and total operator(target) sequence respectively, Kop is the dissociation constant of the protein for the operator, (N_t) is the total concentration of nonspecific sites- generally considered to be 10^{-2} M if each bp represents the start of a new nonspecific site (), K_N is the dissociation constant for nonspecific DNA binding, and (R_t) is the concentration of repressor protein in the nucleus. Then an evolved 6-finger protein with a 10^{-12}M affinity for its target and 10^{-4}M affinity for nonspecific sites, would occupy its target sequence 99% of the time when present at 10nM (5,000 molecules)in the nucleus; $(O)/(O_t) = 10^{-2}$. Application of 10ug/ml or 450nM extracellular protein (MW of 22kD) should be sufficient given no barriers to uptake or overwhelmingly sufficient with active uptake. Evidence from natural model systems suggest that the affinity and dosage estimated in the calculation above for a six-fingered protein above are greatly exaggerated. Trans-activation of the HIV-1 LTR by addition of 100nM extracellular tat protein in the presence of chloroquine to culture medium stimulates trans-activation about 7000-fold compared to controls lacking tat protein(30). Tat binds TAR RNA with an affinity of ~3nM (64). Indeed, trans-activation can be detected when Tat is provided extracellularly at a level as low as 1nM. In a recent report (32)

human lactoferrin has been demonstrated to be taken up by human myelogenous leukemia cells from the culture medium, directed to the nucleus where it binds in a sequence specific fashion to DNA, and activates transcription of a reporter gene containing the lactoferrin specific target sequence. Dose dependence in this case was not well studied but indicated that 10ug/ml or 115nM of iron saturated lactoferrin added to the medium resulted in >100-fold activation of the reporter gene. Lactoferrin binds its 9bp consensus sequence with an affinity of 12nM. The caveat of this comparison is that we are comparing operator occupancy with measured activation levels. Nonetheless, it should be reasonable based on our preliminary data with the 6C7 protein that the ratio $(N_t)/K_N$ for our evolved proteins will be roughly similar to those of Zif268 or tat and lactoferrin. Thus biological efficacy and therefore occupancy of the target site $(O)/(O_t)$ of our proteins should be directly proportional to $K_{op}/(R_t)$ where (R_t) is directly proportional to uptake/nuclear targeting efficiency and extracellular concentration. Thus any increase in affinity of the zinc finger proteins or increase in uptake/nuclear targeting efficiency will dramatically reduce the required extracellular dosage. The issue of the magnitude of (R_t) is only of concern in extracellular delivery, since intracellular expression in the gene therapy application of these proteins could easily meet the production demands of even a low affinity protein. Indeed, intracellular expression of a low affinity (10^{-7} M) three-finger protein by Klug's group at the MRC has demonstrated transcriptional inhibition by blocking pol II passage through a coding sequence of a BCR-ABL fusion oncogene(28).

Tat is believed to bind nonspecifically to cell surfaces with > 10^7 sites per cell and to be internalized by an absorptive endocytosis pathway common to many viruses with charged viral receptors(63). No specific receptor is thought to be involved. Uptake may be dependent on the very basic character, high pI, of Tat since uptake is inhibited by heparin and dextran sulfate. Following uptake, Tat becomes localized in the nucleus. The utility of Tat fusions to deliver heterologous proteins into cells has recently been reported(33). In this report it is estimated that as many as 10^7 molecules of Tat may be internalized in a trypsin insensitive form when provided in the culture medium. Conjugates of Tat peptide fragments with proteins as large as beta-galactosidase (MW of 540kD) have demonstrated internalization into the cytoplasm and nuclear localization. For beta-galactosidase, conjugation at the level of 1 or 2 Tat peptides per tetramer resulted in the internalization of as many as 10^6 molecules per cell. Furthermore this activity was not dependent on cell type. Uptake of beta-galactosidase conjugate was demonstrated with HeLa, COS-1, CHO, H9, NIH 3T3, primary human keratinocytes, and umbilical vein endothelial cells. Further, i.v. injection into mice revealed uptake in heart, liver, spleen, lung, muscle, with little or no activity in the brain or kidney. Within Tat the minimal region required for binding and uptake are residues 47-58 of the basic region, though for some proteins residues 37-72 were most efficient. It is important to note that residues 37-72 bind TAR with an affinity of 84nM while residues 47-58 are insufficient for specific TAR binding(64). For this application, while TAR binding to our tagged proteins would produce a dominant negative inhibitor of Tat, it could sequester the finger proteins and inhibit their action. With human lactoferrin, the region responsible for uptake has not been delineated. Lactoferrin does contain a highly basic region near the N-terminus of the protein which by analogy with the basic region of Tat may serve this role. A minimal human

lactoferrin sequence may be advantageous if the minimal Tat sequence elicits an immune response.

To study uptake and targeting we will construct fusion proteins wherein the high affinity three finger C7 protein is fused at its carboxy terminus, in separate constructs, with 1) residues 37-72 of Tat, 2) residues 47-58 of Tat 2) the nuclear localization signal RRRGL of human angiogenin(65) 3) 100 AA fragments of the human lactoferrin protein, beginning with the N-terminal basic region until a minimal sequence is delineated. Each protein will be tagged as a fusion with the HA peptide epitope to facilitate purification and detection. The two Tat fusions have now been constructed and verified by sequencing. They are currently being expressed and purified using the anti-HA column. Fragments of lactoferrin will be generated by PCR. The expression vector contains the unique cloning sites SpeI, SfiI, and BamHI which facilitate the cloning of inserts in-frame with the fingers and the peptide tag. Retention of DNA binding activity and affinity can be quickly determined using the ELISA methodology described above.

Translocation and targeting will be studied by immunofluorescence and staining (33,66). Cells, HeLa, H9, Jurkat, and the monocytic cell line BF24(), will be incubated (0-18 hrs) with different concentrations of finger proteins. Cells will be washed, fixed, and incubated with anti-HA antibody fluorescently or enzymatically labeled. Functional activity will be assessed with reporter gene assays. Initial characterization will be with variants of the pGL3 luciferase reporter vectors from Promega. Introduction of 5 point mutations in the SV40 promoter region of this vector allows for retention of the Sp1 binding sites and the introduction of two C7 or Zif268 binding sites (GCGTGGGCG) into the complementary strand shown of the sequence given below. Mutations are shown in bold font. Binding to the C7 sites should inhibit Sp1 binding and result in repression. Three Sp1 and the TATA binding regions are also indicated. This plasmid is now under construction. We will study the relative efficacy of repression by binding to promoter elements as compared to targeting coding regions of proteins by introduction of two C7 binding sites into the N-terminus as a fusion with the luciferase coding region. This will append several aa's to the native terminus and is not expected to have an effect on activity.

C7 C7

CCGCCCAGTTCCGCCCACGCTCCGCCCACGCGCTGACTAATTTTTTTTATTTATG

Dose-response curves will be established for the various proteins by transfection mediated by calcium phosphate of reporter plasmids into the Jurkat cell line as recommended by the manufacturer. As internal controls we will use pCH110 from Pharmacia; beta-galactosidase assay, or a SEAP vector, gift of B. Cullen. SEAP can be assayed in a luminometer as well as spectrophotometrically. For a specificity control, the pGL3 vector lacking target sites will be used. Since C7 is the high affinity variant of Zif268 and binds the same target site we will study the effect of the 13-fold increase in affinity of C7 over Zif268 on the dose-response curve. The effect of addition of chloroquine to the medium, demonstrated to enhance the efficacy of Tat, will also be studied. Northern analysis will be used to confirm that repression correlates with decreased specific message. To enhance the ability of these fusions to mediate repression we will also explore the attachment of the recently described human KRAB repressor domain

which can functionally inhibit transcription in a distance and orientation independent fashion(13). Studies have indicated that 42 AA's of this domain is sufficient for complete activity. This fragment will be constructed using oligonucleotides and fused with the finger constructs described above for study. To estimate the biological half-life of the most optimal proteins, they will be incubated with serum for varying amounts of time before being added to the culture medium.

To study the potential of evolved finger proteins to activate gene expression when provided extracellularly, we will fuse activation domains in combination with uptake/targeting sequences to the C7 and Zif268 proteins. For proximal activation we will study the minimal activation domains from human Sp1, AA132-243 and AA340-485, in separate constructs since sequence of activation domains may affect uptake. This will be accomplished using the pGL3 luciferase reporter system described above where the two TATA proximal Sp1 sites are converted to the C7 target sequence. For remote activation we will study fusions of AA286-518 and AA520-550 from the human NF-kappaB activation domain. For this, two C7 binding sites will be inserted into the pGL3 promoter vector upstream of the promoter. The activity of these individual domains have been previously studied and are available from W. Schaffner(12).

To optimize uptake and nuclear targeting we will investigate two strategies. In the first strategy we attempt to capitalize on recent results that filamentous phage displaying RGD sequences as fusions with the gVIII coat protein can be internalized in mammalian cells(67). This provides a direct route to selection for internalization. A collection of available gVIII peptide libraries will be selected for internalization by addition to Jurkat cells followed by incubation, washing and trypsinization to remove surface associated phage. Internalized phage will be released by lysis and allowed to reinfect E. coli. Should they no longer be infective recovery will be made by PCR. Several rounds of such selection will allow for the best sequences from a pool of 10^9 to be studied. Candidate sequences can then be amplified by PCR and fused with the finger constructs above with nuclear localization signals for comparative study. Cell specificity of the motif or lack there of can also be studied. Should this rapid selection fail we will attempt to isolate optimal sequences by screening. A library wherein the alkaline phosphatase gene from the SEAP vector is replaced following the leader sequence with inserts consisting of the C7 protein fused with a nuclear localization signal, a randomized 6 amino acid insert, and an activation domain will be transiently expressed and secreted in HeLa cells. Plating and cocultivation of HeLa cells expressing a stable beta-galactosidase reporter with C7 response elements will allow for transients expressing optimized uptake sequences to be identified. Improvement of known motifs will be screened in a similar way wherein 4 residues within the motif are mutagenized. Prior to introduction into HeLa cells, the libraries will be selected on phage so that only clones which produce functional protein will be screened in the HeLa assay.

Specific Aim 3: Characterize the efficacy of evolved proteins to inhibit HIV-1 transcription and replication using reporter genes, p24 production assays, and integrase assays. Evolved proteins will be:
(i) provided extracellularly as fusions with uptake and targeting sequences.
(ii) intracellularly expressed as envisioned in a gene therapy application.

This phase of study will utilize the evolved fingers constructed above which target HIV-1. We will utilize the optimized conditions and uptake/targeting sequences defined in aim 2. For reporter gene studies we will transfect the plasmid HIV-gpt into the 1G5 cell line (both from the AIDS reagent program). The 1G5 cell line is a Jurkat derivative containing a stable integrated HIV-LTR-Luciferase construct(68). Plasmid HIV-gpt contains a replication defective HXB2 proviral sequences wherein the gp160 coding region has been replaced with the gpt gene. The luciferase gene of the cell line will be responsive to Tat produced from HIV-gpt which provides all the targets for the evolved anti-HIV zinc finger proteins described above. Decrease in luciferase activity dependent on extracellular finger proteins will allow the efficacy of these proteins to inhibit transcription to be determined. Luciferase activity will be correlated with p24 levels determined by ELISA. Transfection controls will be provided as described above. Dose-response will be studied for individual proteins as well as cocktails of evolved zinc finger proteins. To study the efficacy of intracellularly expressed fingers, they will be cloned with nuclear localization signals into the mammalian expression vector pcDNA 3 from Invitrogen. Assays will be as described above with extracellular protein being replaced with transfection of this additional plasmid.

Following the reporter assays are a series of experiments designed to assess breadth and potency in more biologically relevant systems. The effect of evolved fingers on chronically infected H9/HTLVIIIB cells will studied by addition of protein to the medium and by transfection with the finger expression vector. Following 5 days of growth, culture supernatants will be assayed for p24 and compared to control cultures. The pcDNA plasmid contains a neomycin resistance marker which will be used to select stable Jurkat cell lines expressing the evolved fingers. This will also allow us to determine if there are any deleterious effects due to long term expression of the evolved proteins since we can compare cell growth and viability data with controls. Our collaborator, Flossie Wong-Staal has performed similar studies with cell lines expressing her ribozymes and will aid in these studies, see letter. We will then assess the resistance of these cell lines to infection with lab adapted and primary isolates of HIV-1 with p24 production as a readout. Since uptake of extracellular finger proteins may be cell-type dependent we will also examine the function of the fingers in PBMC based assays. In this assay, varying concentrations of evolved finger proteins will be incubated with PBMCs, after which a fixed amount of primary virus (50-100 $TCID_{50}$) will be added. The p24 antigen levels from the finger treated cultures will be compared to control cultures with the non-HIV Zif268 protein. We are currently establishing these assays together with Dennis Burton to evaluate evolved antibody function. Primary viruses to be tested include SI and non-SI clade B viruses, members of each of the clades A to F, and O viruses. Primary viruses will also be used to infect the 1G5 Jurkat cell line which have been incubated with evolved fingers or are transiently expressing them. In this case, in addition to p24 assays, the LTR driven luciferase gene provides a Tat sensitive reporter as described above. Promising proteins and stable cell lines will be provided to John Moore of the Aaron Diamond Institute for assays with additional primary clinical isolates and to Joseph Sodroski of the Dana-Farber for additional reporter and viral production assays, see attached letters.

As a more exploratory investigation, we will assess the ability of evolved finger proteins to inhibit integration of HIV-1. Since integration of viral DNA into the host genome is an early

essential step in the lifecycle of the virus, inhibition of this process is attractive. Unfortunately, in these studies we are not guided by preliminary results which would suggest attractive target sequences which upon binding by an evolved finger would inhibit integration perhaps by displacing the integrase. These assays are discovery driven and simply consist of screening our proteins in an *in vitro* assay for inhibition of integration. Early work by Varmus, Bishop, and colleagues have demonstrated that viral DNA in the active nucleoprotein complex is accessible to nucleases and thus should be accessible to DNA binding proteins(69). Assays(70,71) will be performed in collaboration with Frederic Bushman of the Salk Institute, see attached letter. Since recent studies suggest that virus expression is directly involved in $CD4^+$ cell destruction(1,2), given time we will explore the possibility of inducing rapid cell death by selective hyperactivation of HIV-1 expression. For these studies evolved 6-finger proteins with activation domains described above will be incubated with latent and chronically infected cells (AIDS Reagent Program). Cell viability will then be studied with appropriate control proteins. In combination with drugs which inhibit *de novo* infection, such a strategy could eliminate long-lived latent and chronically infected cell populations which may be important in the disease.

Specific Aim 4: Study the structural basis for sequence-specific recognition of DNA by evolved zinc finger proteins in collaboration with Dr. Peter Wright.

These studies will focus on the novel polydactyl proteins generated to inhibit HIV-1. Joel Gottesfeld will join these studies to correlate the sequences and structures of evolved fingers with their functional properties of affinity and sequence-specific recognition. Though a simple recognition code for these proteins has been all but discounted, structural study will provide important insights into the properties of these proteins and perhaps shed light on their unexpectedly high abundance in our own genomes. The HIV-1 target sequences are rich in sequences which have not been structurally observed in complex with zinc finger proteins. For example, the PBS region specific protein will allow for a first look at recognition of an AAA stretch and may provide insight into why the C7 finger recognizes its GCG site with higher affinity and specificity than the natural Zif268 finger. The 6-finger proteins promise to present a topologically exciting structure which involves the protein wrapping around the DNA target for nearly two full turns of the helix. We will provide Peter Wright with required proteins for his NMR experiments. Dr. Wright is currently funded to investigate the structural aspects of zinc finger DNA interactions, see attached letters.

Time Table: Year 1: Aims 1(i and ii) and 2
Year 2: Aims 1(i and ii), 2,3, and 4
Year 3: Aims 1(ii), 3, and 4

5. Human Subjects: None

6. Vertebrate Animals: None

7. Collaborators: Frederic D. Bushman, Ph.D.
John P. Moore, Ph.D.
Flossie Wong-Staal, Ph.D.
Joel M. Gottesfeld, Ph.D.
Joseph G. Sodroski, M.D.
Peter E. Wright, Ph.D.

8. Consortium: None

9. Literature Cited: 5 pages

The Salk Institute for Biological Studies

Frederic D. Bushman, Ph.D.
Infectious Disease Laboratory

April 19, 1995

Carlos F. Barbas III, Ph.D.
Assistant Member
The Scripps Research Institute
10666 North Torrey Pines Road
La Jolla, CA 92037

Dear Carlos:

This letter is to confirm my interest in collaborating on the development of evolved zinc fingers proteins as inhibitors of HIV-1 integration. We have established rapid *in vitro* assays for HIV DNA integration and would be very interested in examining the ability of your zinc finger proteins to inhibit the integrase. We are currently involved in characterizing the integration complex and will work with you to effectively target the DNA for inhibition of complex formation. Zinc Fingers targeted to the LTR's as well as more central coding regions should be examined. In the long run, I am also interested in using your zinc finger proteins to direct integration to specific sites in the genome. This may allow gene therapy to be practiced in a more controlled fashion.

Sincerely,

Frederic Bushman

Post Office Box 85800 • San Diego, CA • 92186-5800 • (619) 453-4100 • Fax (619) 552-8285

The Aaron Diamond AIDS Research Center
For the City of New York

March 28, 1995

Dr. Carlos Barbas, Ph.d.
The Scripps Research Institute
Department of Molecular Biology
10666 North Torrey Pines Road
La Jolla, CA 92037

Dear Carlos;

It is with great pleasure that I confirm that we will be delighted to collaborate with you on your project to prepare zinc finger proteins as antiviral agents aimed to inhibit HIV transcription and replication. In particular, the facilities of my laboratory will be available to test your reagents for efficacy against lab strain and primary isolates of HIV-1.

Yours is an imaginative proposal at the cutting edge of current technology; but you have the technical skills and the drive to make a project like this succeed.

There has been, over the past two years, a constructive and productive collaboration between you and your colleagues at the Scripps Research Institute and us at Aaron Diamond. This will continue, and it is something I look forward to very much. The recent move of James Binley from your Center to ours only facilitate these interactions.

I wish you every success with your proposal; it deserves to succeed.

Sincerely,

John P. Moore, Ph. D.
Staff Investigator
Aaron Diamond AIDS Research Center
Associate Professor of Medicine and Microbiology
New York University School of Medicine

JPM:jtg

455 1ST AVENUE, 7TH FLOOR ♦ NEW YORK, NEW YORK 10016
TELEPHONE: (212) 725-0018 FAX (212) 725-1126

UNIVERSITY OF CALIFORNIA, SAN DIEGO UCSD

BERKELEY • DAVIS • IRVINE • LOS ANGELES • RIVERSIDE • SAN DIEGO • SAN FRANCISCO SANTA BARBARA • SANTA CRUZ

FLOSSIE WONG-STAAL, PH.D.
FLORENCE SEELEY RIFORD PROFESSOR
OF ACQUIRED IMMUNE DEFICIENCY
SYNDROME RESEARCH

DEPARTMENT OF MEDICINE
9500 GILMAN DRIVE, 0665
LA JOLLA, CALIFORNIA 92093-0665
(619) 534-7957
(619) 534-7743 FAX

April 10, 1995

Carlos F. Barbas III, Ph.D.
Assistant Member
The Scripps Research Institute
10666 North Torrey Pines Road
La Jolla, CA 92037

Dear Carlos:

I am excited about your proposed project to inhibit HIV-1 transcription and replication with zinc finger proteins. Your recent PNAS publication certainly demonstrates that using your techniques you can select zinc fingers to bind new target sequences. These could be utilized in ways analogous and perhaps complimentary to my catalytic RNAs. Should your zinc finger project develop along the gene therapy route, I would be happy to collaborate with you to develop this strategy with your proteins.

Sincerely yours,

Flossie Wong-Staal, Ph.D.
Professor of Medicine and Biology

FWS/es

THE SCRIPPS RESEARCH INSTITUTE

10666 North Torrey Pines Road
La Jolla, CA 92037
Telephone 619.455.9100

Joel M. Gottesfeld, Ph.D.
Member, Department of Molecular Biology

Direct Line: 619.554.8461
Fax: 619.554.6665

April 13, 1995

Dr. Carlos Barbas
Department of Molecular Biology
The Scripps Research Institute

Dear Carlos,

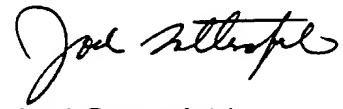
I am writing this letter to confirm my willingness to continue our very productive collaboration on the design and selection of zinc finger proteins that will target particular gene sequences. As you are aware, my laboratory has identified the zinc finger domains of the 5S gene-specific transcription factor TFIIIA that are responsible for high affinity interactions with both the 5S RNA gene and with 5S RNA itself. We have also determined both the DNA binding sites and RNA binding sites for these sets of zinc fingers. We are continuing our studies on the mechanisms of DNA and RNA recognition by these zinc finger polypeptides. Structural studies in Peter Wright's laboratory on these protein-DNA and RNA complexes will provide important insights into the differential mechanisms of nucleic acid recognition by this fascinating protein.

As you are aware, we have examined whether a zinc finger polypeptide can block the transcription of an active gene harboring the binding site for that polypeptide. To this end, we carried out a series of *in vitro* transcription experiments in which we monitored transcription of various DNA templates containing the binding site for TFIIIA zinc fingers 1-3 (zf1-3) in the presence or absence of this polypeptide. We designed plasmids harboring the zf 1-3 binding site downstream from the bacteriophage T7 promoter and upstream from a tRNA gene sequence. Similar amounts of zf 1-3 polypeptide were required to observe DNA binding and inhibition of transcription. The graphs on the following pages show quantitation of these data as plots of percent inhibition of transcription versus percent of DNA molecules bound by the zinc finger protein (as determined by gel mobility shift DNA binding assays). These results show that zinc finger polypeptides, whether delivered as the polypeptide itself or in the form of its coding sequence, may be useful in modulating gene transcription in animal or plant cells. Modulation of gene activity will allow a new therapeutic approach to a variety of disease states including various viral infections and cancer.

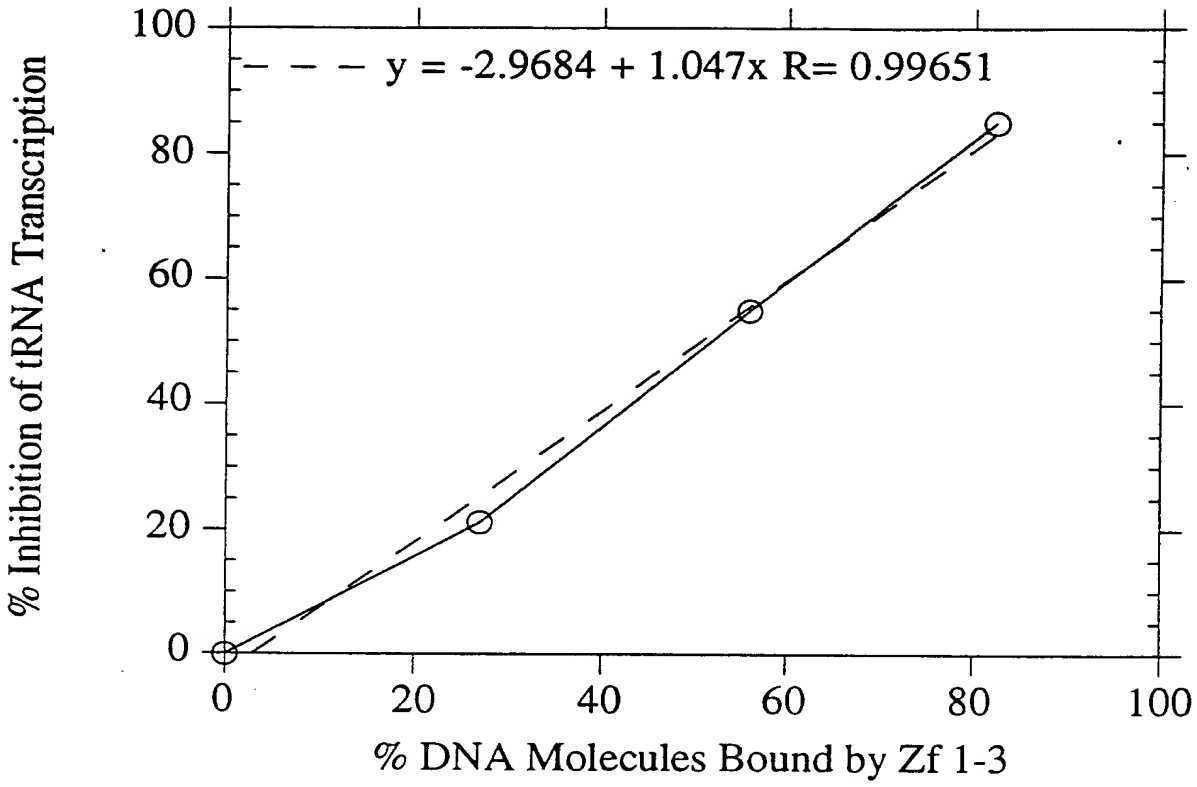
My laboratory will be pleased to provide you with any materials you may need for your research, such as recombinant zinc finger proteins, expression plasmids and DNA or RNA probes. We will also be happy to provide any assistance in protein expression/purification and binding assays.

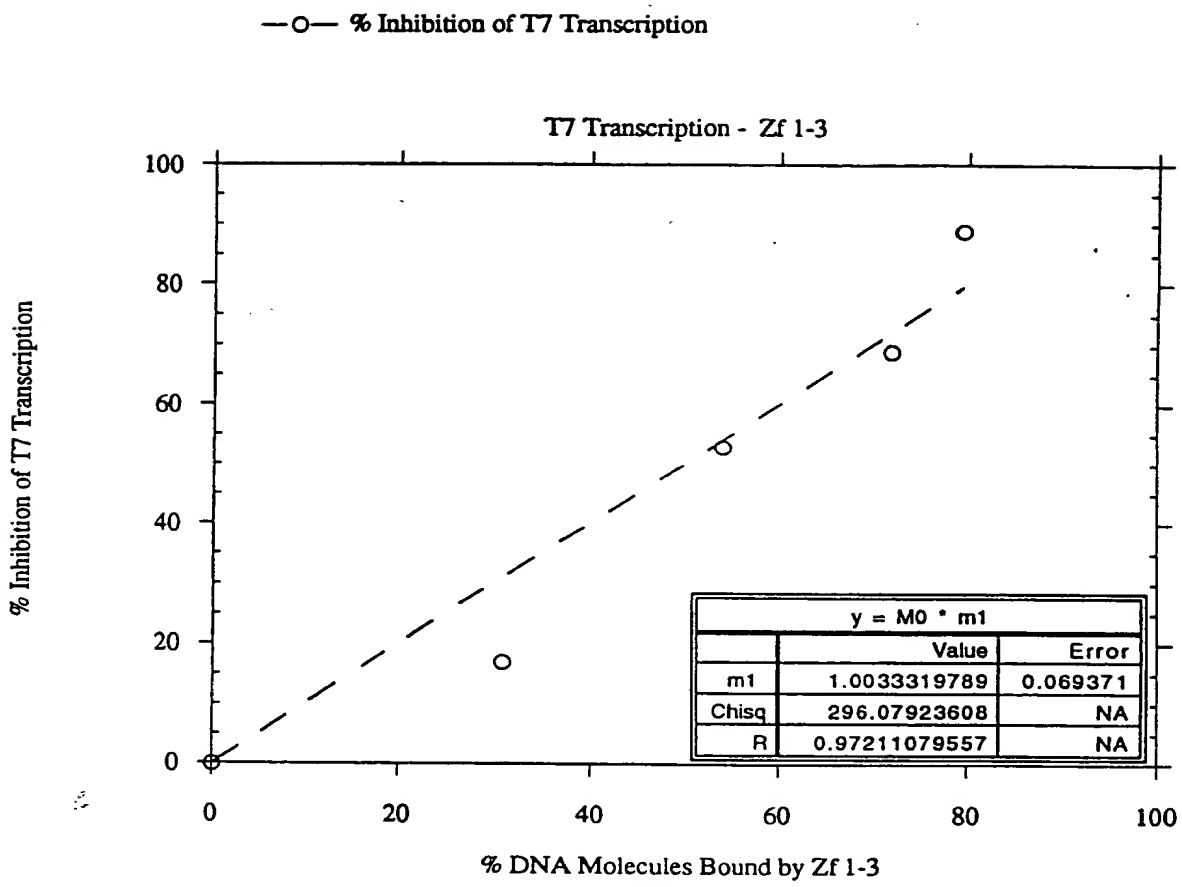
I'll look forward to many more years of productive collaboration with you and members of your group.

Sincerely yours,



Joel Gottesfeld


y = -2.9684 + 1.047x R= 0.99651
% Inhibition of tRNA Transcription
% DNA Molecules Bound by Zf 1-3
0
20
40
60
80
100


—o— % Inhibition of T7 Transcription
T7 Transcription - Zf 1-3
% Inhibition of T7 Transcription
% DNA Molecules Bound by Zf 1-3
0
20
40
60
80
100
y = M0 * m1
Value
Error
m1
1.0033319789
0.069371
Chisq
296.07923608
NA
R
0.97211079557
NA

DANA-FARBER
CANCER INSTITUTE

44 Binney Street, Boston, MA 02115

JOSEPH G. SODROSKI, M.D.
Chief, Division of Human Retrovirology
Tel. 617-632-3371
Fax 617-632-4338

HARVARD MEDICAL SCHOOL
Associate Professor of Pathology

April 6, 1995

Carlos F. Barbas III, Ph.D.
Department of Molecular Biology
The Scripps Research Institute
10666 North Torrey Pines Road
La Jolla, CA 92037

Dear Carlos:

I would be happy to collaborate with you on the development of engineered zinc finger proteins targeting the HIV-1 LTR. As you know, we have developed assay systems that provide quantitative assessments of HIV-1 replication. We would be interested in examining the effects of any reagents that you develop on HIV-1 replication in these assays. Since we also have HIV-1 vectors that express marker genes under the control of promoters other than the HIV-1 LTR, we can examine the specificity of any observed effects as well. I look forward to an interesting collaboration.

Sincerely yours,

Joseph G. Sodroski, M.D.
Chief, Division of Human Retrovirology
Dana-Farber Cancer Institute

JGS/jw

THE SCRIPPS RESEARCH INSTITUTE

10666 NORTH TORREY PINES ROAD, LA JOLLA, CALIFORNIA 92037

Peter E. Wright, Ph.D.
Cecil H. and Ida M. Green Investigator
Member and Chairman
Department of Molecular Biology

Telephone: 619/554-9721
Facsimile: 619/554-9822

April 1, 1995

Dr. Carlos Barbas, III
Department of Molecular Biology
The Scripps Research Institute
10666 North Torrey Pines Road
La Jolla, CA 92037

Dear Carlos:

I would be delighted to continue our collaboration on the structural characterization of designed zinc finger proteins. As you know, we have an intense interest in elucidating the structural basis for sequence specific DNA and RNA recognition by TFIIIA and related zinc finger proteins and have recently determined low resolution NMR structures for three zinc fingers bound to DNA. We look forward to working with you to determine structures of your novel zinc fingers selected for binding predetermined target sequences. I am extremely excited about this project and believe that we will obtain important new insights into nucleic acid targeting by zinc fingers and the mechanisms of sequence specific recognition.

I wish you good luck with your grant application.

With best wishes.

Sincerely yours,

P. E. Wright

Peter E. Wright

PEW/mcs

Barbas III, C.F.

References:

1. Wei, X., Ghosh, S.K., Taylor, M.E., Johnson, V.A., Emini, E.A., Deutsch, P., Lifson, J.D., Bonhoeffer, S., Nowak, M.A., Hahn, B.H., Saag, M.S., and Shaw, G.M. (1995) Viral dynamics in human immunodeficiency virus type 1 infection. *Nature* **373**:117-122.
2. Ho, D.D., Neumann, A.U., Perelson, A.S., Chen, W., Leonard, J.M., and Markowitz, M. (1995) Rapid turnover of plasma virions and CD4 lymphocytes in HIV-1 infection. *Nature* **373**:123-126.
3. Matsukura, M., Zong, G., Shinozuka, K., Robert-Guroff, M., Shimada, T., Stein, C.A., Mitsuya, H., Wong-Staal, F., Cohen, J.S., and Broder, S. (1989) Regulation of viral expression of human immunodeficiency virus *in vitro* by an antisense phosphorothioate oligodeoxynucleotide against *rev (art/trs)* in chronically infected cells. *Proc. Natl. Acad. Sci. USA* **86**:4244-4248.
4. Stein, C.A., and Cheng, Y.-C. (1993) Antisense Oligonucleotides as Therapeutic Agents - Is the Bullet Really Magical? *Science* **261**:1004-1012.
5. Baltimore, D. (1988) Intracellular Immunization. *Nature* **335**: 395-396.
6. Yu, M., Ojwang, J., Yamada, O., Hampel, A., Rapapport, J., Looney, D., and Wong-Staal, F. (1993) A hairpin ribozyme inhibits expression of diverse strains of human immunodeficiency virus type 1. *Proc. Natl. Acad. Sci. USA* **90**:6340-6344.
7. Yu, M., Leavit, M.C., Maruya, M., Yamada, O., Young, D., Ho, A.D., and Wong-Staal, F. (1995) Intracellular immunization of human fetal cord blood stem/progenitor cells with a ribozyme against human immunodeficiency virus type 1. *Proc. Natl. Acad. Sci. USA* **92**:699-703.
8. Sullenger, B.A., Gallardo, H.F., Ungers, G.E., and Gilboa, E. Overexpression of TAR Sequences Renders Cells Resistant to HIV Replication. (1990) *Cell* **63**: 601-608.
9. Ptashne, M. A Genetic Switch Phage l and Higher Organisms. Cell Press & Blackwell Scientific Publications, Cambridge, MA, 1992.
10. Herschbach, B.M., Arnaud, M.B., and Johnson, A.D. (1994) Transcriptional repression directed by the yeast a2 protein *in vitro*. *Nature* **370**:309-311.
11. Kashanchi, F., Duvall, J.F., Lindholm, P.F., Radonovich, M.F., and Brady, J.N. (1993) Sequences Downstream of the RNA Initiation Site Regulate Human T-Cell Lymphotropic Virus Type I Basal Gene Expression. *J. Virol.* **67**:2894-2902.
12. Seipel, K., Georgiev, O., and Schaffner, W. (1992) Different activation domains stimulate transcription from remote ('enhancer') and proximal ('promoter') positions. *EMBO J.* **11**:4961-4968.
13. Pengue, G., Calabro, V., Bartoli, P.C., Pagliuca, A., and Lania, L. Repression of transcriptional activity at a distance by the evolutionarily conserved KRAB domain present in a subfamily of zinc finger proteins. (1994) *Nucleic Acids Research*, **22**: 2905-2914.
14. Rhodes, D., and Klug, A. (1993) Zinc Fingers. Scientific American, pgs. 56-65.
15. Jacobs, G.H. (1992) Determination of the base recognition positions of zinc fingers from sequence analysis. *EMBO J.* **11**:4507-4517.
16. Lee, M.S., Gippert, G.P., Soman, K.V., Case, D.A., and Wright, P.E. (1989) Three-Dimensional Solution Structure of a Single Zinc Finger DNA-Binding Domain. *Science* **245**:565-680.

17. Desjarlais, J.R., and Berg, J.M. (1993) Use of a zinc-finger consensus sequence framework and specificity rules to design specific DNA binding proteins. *Proc. Natl. Acad. Sci. USA* **90**:2256-2260.
18. Pavletich, N.P., and Pabo, C.O. (1991) Zinc Finger-DNA Recognition: Crystal Structure of Zif268-DNA Complex at 2.1 Å. *Science* **252**:809-817.
19. Pavletich, N.P., and Pabo, C.O. (1993) Crystal Structure of a Five-Finger GLI-DNA Complex: New Perspectives on Zinc Fingers. *Science* **261**:1701-1707.
20. Fairall, L., Schwabe, J.W.R., Chapman, L., Finch, J.T., and Rhodes, D. (1993) The crystal structure of a two zinc-finger peptide reveals an extension to the rules for zinc-finger/DNA recognition. *Nature* **366**:483-487.
21. Clemens, K.R., Wolf, V., McBryant, S.J., Zhang, P., Liao, X., Wright, P.E., Gottesfeld, J.M. (1993) Molecular Basis for Specific Recognition of Both RNA and DNA by a Zinc Finger Protein. *Science* **260**:530-533.
22. Shi, Y., and Berg, J.M. (1995) Specific DNA-RNA Hybrid Binding by Zinc Finger Proteins. *Science* **268**:282-284.
23. Wu, H., Yang, W.-P., and Barbas III, C.F. (1995) Building Zinc Fingers by selection: Towards a therapeutic application. *Proc. Natl. Acad. Sci. USA* **92**:344-348.
24. Yang, W.-P., Wu, H., and Barbas III, C.F. (1995) Surface Plasmon Resonance based Kinetic Studies of Zinc Finger-DNA Interactions. *Methods in Immunology*, in press.
25. Rebar, E.J., and Pabo, C.O. (1994) Zinc Finger Phage: Affinity Selection of Fingers with New DNA-Binding Specificities. *Science* **263:** 671-673.
26. Jamieson, A.C., Kim, S.-H., and Wells, J.A. (1994) *In Vitro* Selection of Zinc Fingers with Altered DNA-Binding Specificity. *Biochemistry* **33**:5689-5695.
27. Choo, Y., and Klug, A. (1994) Toward a code for the interactions of zinc fingers with DNA: Selection of randomized fingers displayed on phage. *Proc. Natl. Acad. Sci. USA* **91**:11163-11167.
28. Choo, Y., Sanchez-Garcia, I., and Klug, A. (1994) In vivo repression by a site-specific DNA-binding protein designed against an oncoginc sequence. *Nature* **372**:642-645.
29. Pomerantz, J.L., Sharp, P.A., Pabo, C.O. (1995) Structure-Based Design of Transcription Factors. *Science* **267**:93-96.
30. Frankel, A.D., and Pabo, C.O. (1988) Cellular Uptake of the Tat Protein from Human Immunodeficiency Virus. *Cell* **55**:1189-1193.
31. Lindholm, P.F., Marriott, S.J., Gitlin, S.D., Bohan, C.A. & Brady, J.N. (1990) Induction of nuclear NF-kappaB DNA Binding Activity after exposure of lymphoid cells to soluble Tax-1 Protein *New Biologist* **2:** 1034-1043.
32. He, J., and Furmanski, P. (1995) Sequence specificity and transcriptional activation in the binding of lactoferrin to DNA. *Nature* **373**:721-724.
33. Fawell, S., Seery, J., Daikh, Y., Moore, C., Chen, L.L., Pepinsky, B., and Barsoum, J. (1994) Tat-mediated delivery of heterologous proteins into cells. *Proc. Natl. Acad. Sci. USA* **91**:664-668.
34. Myers, G., Korber, B., Berzofsky, J.A., Smith, R.F., and Pavlakis, G.N. (1994) Human Retroviruses and AIDS (Theoretical Biology and Biophysics, Los Alamos, NM)

35. Kang, A.S., Barbas III, C.F., Janda, K.D., Benkovic, S.J., and Lerner, R.A. (1991) Linkage of recognition and replication functions by assembling combinatorial antibody Fab libraries along phage surfaces. *Proc. Natl. Acad. Sci. USA* **88**:4363-4366.
36. Barbas III, C.F., Kang, A.S., Lerner, R.A., and Benkovic, S.J. (1991) Assembly of combinatorial antibody libraries on phage surfaces: The gene III site. *Proc. Natl. Acad. Sci. USA* **88**:7978-7982.
37. Burton, D.R., and Barbas III, C.F. (1994) Human antibodies from Combinatorial Libraries. *Advances in Immunology* **57**:191-280.
38. Barbas III, C.F., Bain, J.D., Hoekstra, D.M., and Lerner, R.A. (1992) Semi-synthetic combinatorial antibody libraries: A chemical solution to the diversity problem. *Proc. Natl. Acad. Sci. USA* **89**:4457-4461.
39. Janda, K.D., Lo, C.-H.L., Ti, T., Barbas III, C.F.,Wirsching, P., and Lerner, R.A. (1994) Direct Selection for a Catalytic Mechanism from Combinatorial Antibody Libraries. *Proc. Natl. Acad. Sci. USA* **91**:2532-2536.
40. Barbas III, C.F., Languino, L.R., and Smith, J.W. (1993) High Affinity Self-Reactive Human Antibodies by Design and Selection: Targeting the Integrin Ligand Binding Site. *Proc. Natl. Acad. Sci. USA* **90**:10003-10007.
41. Smith, J.W., Hu, D., Satterthwait, A.C., Pinz-Sweeney, S., and Barbas III, C.F. (1994) Binding Synthetic Antibodies as Adhesive Ligands for Integrins. *J. Biol. Chem.* **269**:32788-32795.
42. Barbas III, C.F., Crowe, J.E., Cababa, D., Jones, T.M., Zebedee, S.L., Murphy, B.R., Chanock, R.M., and Burton, D.R. (1992) Human Monoclonal Fab Fragments derived from a combinatorial library bind to respiratory syncytial virus F glycoprotein and neutralize infectivity. *Proc. Natl. Acad. Sci. USA* **89**:10164-10168.
43. Crowe Jr., J.E., Murphy, B.R., Chanock, R.M., Williamson, R.A., Barbas III, C.F., and Burton, D.R. (1994) Recombinant human respiratory syncytial virus (RSV) monoclonal antibody Fab is effective therapeutically when introduced directly into the lungs of RSV-infected mice. *Proc. Natl. Acad. Sci. USA* **91**:2133-2137.
44. Barbas III, C.F., Collet, T.A., Roben, P., Binley, J., Amberg, W., Hoekstra, D., Cababa, D., Jones, T.M., Williamson, R.A., Pilkington, G.R., Haigwoods, N.L., Satterthwait, A.C., Sanz, I., and Burton, D.R. (1993) Molecular profile of an antibody response to HIV-1 as probed by combinatorial libraries. *J. Mol. Biol.* **230**:812-823.
45. Barbas III, C.F., Bjorling, E., Chiodi, F., Dunlop, N., Cababa, D., Jones, T.M., Zebedee, S.L., Persson, M.A.A., Nara, P.L., Norrby, E., and Burton, D.R. (1992) Recombinant Human Fab fragments neutralize human type 1 immunodeficiency virus *in vitro*. *Proc. Natl. Acad. Sci. USA* **89**:9339-9343.
46. Burton, D.R., Pyati, J., Koduri, R., Sharp, S.J., Thornton, G.B., Parren, P.W.H.I., Sawyer, L.S.W., Hendry, M.R., and Barbas III, C.F. (1994) Efficient Neutralization of Primary Isolates of HIV-1 by a Recombinant Human Monoclonal Antibody. *Science* **266**:1024-1026.
47. Barbas III, C.F., Hu, D., Dunlop, N., Sawyer, L., Cababa, D., Hendry, R.M., Nara, P.L., and Burton, D.R. (1994) In vitro evolution of a neutralizing human antibody to HIV-1 to enhance affinity and broaden strain cross-reactivity. *Proc. Natl. Acad. Sci. USA* **91**:3809-3813.

48. Yang, W.-P., Green, K., Pinz-Sweeney, S., Briones, A.T., Burton, D.R., and Barbas III, C.F. (1995) CDR Walking Mutagenesis for the Affinity Maturation of a Potent Human anti-HIV-1 Antibody into the Picomolar Range. *J. Mol. Biol.*, submitted.
49. Barbas III, C.F., and Lerner, R.A. (1991) Combinatorial immunoglobulin libraries on the surface of phage (Phabs): Rapid selection of antigen specific Fabs. *Methods: Companion Methods Enzymol.* **2:**119-124.
50. Rauscher III, F.J., Morris, J.F., Tournay, O.E., Cook, D.M., and Curran, T. (1990) Binding of the Wilms' Tumor Locus Zinc Finger Protein to the EGR-1 Consensus Sequence. *Science* **250:**1259-1262.
51. Barbas III, C.F. (1993) Recent advances in phage display. *Curr. Opinion in Biochem.* **4:**526-530.
52. Huang, L., Sera, T., and Schultz, P.G. (1994) A permutational approach toward protein-DNA recognition. *Proc. Natl. Acad. Sci. USA* **91:**3969-3973.
53. Verdin, E. (1991) DNase I-Hypersensitive Sites Are Associated with Both Long Terminal Repeats and with the Intragenic Enhancer of Integrated Human Immunodeficiency Virus Type 1. *J. Virol.* **65:**6790-6799.
54. Kharroubi, A.E., and Verdin, E. (1994) Protein-DNA Interactions within DNase I-hypersensitive Sites Located Downstream of the HIV-1 Promoter. *J.Biol. Chem.* **269:**19916-19924.
55. Vaishnav, Y.N., and Wong-Staal, F. (1991) The Biochemistry of AIDS. *Annu. Rev. Biochem.* **60:**577-630.
56. Rhim, H., Park, J., and Morrow, C.D. (1991) Deletions in the $tRNA^{Lys}$ Primer-Binding Site of Human Immunodeficiency Virus Type 1 Identify Essential Regions for Reverse Transcription. *J. Virol.* **65:**4555-4564.
57. Nagashunmugam, T., Velpandi, A., Goldsmith, C.S., Zaki, S.R., Kalyanaraman, V.S., and Srinivasan, A. (1992) Mutation in the primer binding site of the type 1 human immunodeficiency virus genome affects virus production and infectivity. *Proc. Natl. Acad. Sci. USA* **89:**4114-4118.
58. Jones, K.A. (1993) Tat and the HIV-1 promoter. *Curr. Opinion in Cell Biol.* **5:**461-468.
59. Kamine, J., and Chinnadurai, G. (1992) Synergistic Activation of the Human Immunodeficiency Virus Type 1 Promoter by the Viral Tat Protein and Cellular Transcription Factor Sp1. *J. Virol.* **66:**3932-3936.
60. Zeichner, S.L., Kim, J.Y.H., and Alwine, J.C. (1991) Linker-Scanning Mutational Analysis of the Transciptional Activity of the Human Immunodeficiency Virus Type 1 Long Terminal Repeat. *J. Virol.* **65:**2436-2444.
61. Romaniuk, P.J. (1990) Characterization of the Equilibrium Binding of Xenopus Transcription Factor IIIA to the 5 S RNA Gene. *J. Biol. Chem.* **265:**17593-17600.
62. Blackwell, K.T., and Weintraub, H. Differences and Similarities in the DNA-Binding Preferences of MyoD and E2A Protein Complexes Revealed by Binding Site Selection (1990) *Science* **250**: 1104-1110.
63. Mann, D.A., and Frankel, A.D. (1991) Endocytosis and targeting of exogenous HIV-1 Tat protein. *EMBO J.* **10:**1733-1739.

64. Churchill, M.J., Lamont, C., Hamy, F., Dingwall, C., Green, S.M., Lowe, A.D., Jonathan, P., Butler, G., Gait, M.J., and Karn, J. (1993) High Affinity Binding of TAR RNA by the Human Immunodeficiency Virus Type-1 *tat* Protein Requires Base-pairs in the RNA Stem and Amino Acid Residues Flanking the Basic Region. *J. Mol. Biol.* **230**:90-110.
65. Moroianu, J., and Riordan, J.F. (1994) Identification of the Nucleolar Targeting Signal of Human Angiogenin. *Biochem & Biophy. Research Comm.* **203**:1765-1772.
66. Ensoli, B., Buonaguro, L., Barillari, G., Fiorelli, V., Gendelman, R., Morgan, R.A., Wingfield, P., and Gallo, R.C. (1993) Release, Uptake, and Effects of Extracellular Human Immunodeficiency Virus Type 1 Tat Protein on Cell Growth and Viral Transactivation. *J. Virol.* **67**:277-287.
67. Hart, S.L., Knight, A.M., Harbottle, R.P., Mistry, A., Hunger, H.-D., Cutler, D.F., Williamson, R., and Coutelle, C. (1994) Cell Binding and Internalization by Filamentous Phage Displaying a Cyclic Arg-Gly-Asp-containing Peptide. *J. Biol. Chem.* **269**:12468-12474.
68. Aguilar-Cordova, E., Chinen, J., Donehower, L., Lewis, D.E., and Belmont, J.W. (1994) A Sensitive Reporter Cell Line for HIV-1 *tat* Activity, HIV-1 Inhibitors, and T Cell Activation Effects. *AIDS Research and Human Retroviruses* **10**:295-301.
69. Bowerman, B., Brown, P.O., Bishop, J.M., and Varmus, H.E. (1989) A nucleoprotein complex mediates the integration of retroviral DNA. *Genes and Devel.* **3**:469-478.
70. Craigie, R., Mizuuchi, K., Bushman, F.D., and Engelman, A. (1991) A rapid *in vitro* assay for HIV DNA integration. *Nucleic Acids Research* **19**:2729-2734.
71. Bushman, F.D., and Craigie, R. (1991) Activities of human immunodeficiency virus (HIV) integration protein *in vitro*: Specific cleavage and integration of HIV DNA. *Proc. Natl. Acad. Sci. USA* **88**:1339-1343.

CHECKLIST

TYPE OF APPLICATION

[X] NEW application. *(This application is being submitted to the PHS for the first time.)*

[ ] REVISION of application number: _____
(This application replaces a prior unfunded version of a new, competing continuation, or supplemental application.)

[ ] COMPETING CONTINUATION of grant number: _____
(This application is to extend a funded grant beyond its current project period.)

[ ] SUPPLEMENT to grant number: _____
(This application is for additional funds to supplement a currently funded grant.)

[ ] CHANGE of principal investigator/program director.
Name of former principal investigator/program director: _____

[ ] FOREIGN application, city and country of birth and present citizenship of principal investigator/program director. *(This information is required by the U.S. Department of State.)* _____

1. **ASSURANCES/CERTIFICATIONS**

The following assurances/certifications are made by checking the appropriate boxes and are verified by the signature of the OFFICIAL SIGNING FOR APPLICANT ORGANIZATION on the FACE PAGE of the application. Descriptions of individual assurances/certifications begin on page 24 of Specific Instructions.

a. Human Subjects *(Complete Item 4 on the Face Page)*
[ ] Full IRB Review [ ] Expedited Review

b. Vertebrate Animals *(Complete Item 5 on the Face Page)*

c. Inventions and Patents *(Competing Continuation Application Only—Complete Item 10 on the Face Page)*

d. Debarment and Suspension [X] No [ ] Yes *(Attach explanation)*

e. Drug-Free Workplace *(Applicable only to new or revised applications being submitted to the PHS for the first proposed project period, Type 1)*
[X] Yes [ ] No *(Attach explanation)*

f. Lobbying

With Federal appropriated funds [X] No

With other than Federal appropriated funds [X] No [ ] Yes
(If "Yes," see page 29 and attach Standard Form LLL, "Disclosure of Lobbying Activities," to the application behind the second page of the Checklist.)

g. Delinquent Federal Debt [X] No [ ] Yes *(Attach explanation)*

h. Misconduct in Science (Form PHS 6315) [X] Filed [ ] Not Filed
If filed, date of Initial Assurance or latest Annual Report 1/27/95

i. Civil Rights Form HHS 441	j. Handicapped Individuals Form HHS 641	k. Sex Discrimination Form HHS 639-A	l. Age Discrimination Form HHS 680
[X] Filed	[X] Filed	[X] Filed	[X] Filed
[ ] Not Filed	[ ] Not Filed	[ ] Not Filed	[ ] Not Filed

CHECKLIST *(Continued)*

2. PROGRAM INCOME *(See instructions, page 32.)*

All applications must indicate (Yes or No) whether program income is anticipated during the period(s) for which grant support is requested.

[X] No [ ] Yes If "Yes," use the format below to reflect the amount and source(s) of anticipated program income.

Budget Period	Anticipated Amount	Source(s)

3. INDIRECT COSTS

Indicate the applicant organization's most recent indirect cost rate established with the appropriate DHHS Regional Office, or, in the case of forprofit organizations, the rate established with the appropriate PHS Agency Cost Advisory Office. If the applicant organization is in the process of initially developing or renegotiating a rate, or has established a rate with another Federal agency, it should, immediately upon notification that an award will be made, develop a tentative indirect cost rate proposal. This is to be based on its most recently completed fiscal year in accordance with the principles set forth in the pertinent *DHHS Guide for Establishing Indirect Cost Rates*, and submitted to the appropriate DHHS Regional Office or PHS Agency Cost Advisory Office. Indirect costs will ***not*** be paid on foreign grants, construction grants, grants to Federal organizations, and grants to individuals, and usually not on conference grants. Follow any additional instructions provided for Research Career Development Awards, Institutional National Research Service Awards, and the specialized grant applications listed on page 6.

[X] DHHS Agreement dated: 7/29/94 [ ] No Indirect Costs Requested.

[ ] DHHS Agreement being negotiated with _____ Regional Office.

[ ] No DHHS Agreement, but rate established with _____ Date _____

CALCULATION*

(The entire grant application, including the Checklist, will be reproduced and provided to peer reviewers as CONFIDENTIAL information. Supplying the following information on indirect costs is OPTIONAL for forprofit organizations.)

a. Initial budget period:
Amount of base $ 254,427 x Rate applied 70 % = Indirect costs (1) $ 178,099

b. Entire proposed project period:
Amount of base $ 798,463 x Rate applied 70 % = Indirect costs (2) $ 558,924

(1) Add to total direct costs from form page 4 and enter new total on FACE PAGE, Item 7b.
(2) Add to total direct costs from form page 5 and enter new total on FACE PAGE, Item 8b.

*Check appropriate box(es):

[ ] Salary and wages base [X] Modified total direct costs base [ ] Other base (Explain below)

[ ] Off-site, other special rate, or more than one rate involved (Explain below)

Explanation *(Attach separate sheet, if necessary.):*